

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant: SOSHIRO KUZUNUKI ET AL.

Serial No.: NOT YET ASSIGNED

Filed: FEBRUARY 27, 2002

Title: SPEECH INPUT SYSTEM, SPEECH PORTAL SERVER, AND  
SPEECH INPUT TERMINAL

#5  
TLR  
8/5/02



CLAIM FOR PRIORITY UNDER 35 U.S.C. §119

**Box PATENT APPLICATION**  
Commissioner for Patents  
Washington, D.C. 20231

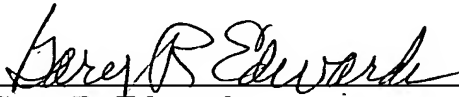
February 27, 2002

Sir:

The benefit of the filing date of prior foreign application No. 2001-306787, filed in Japan on 02 October 2001, is hereby requested and the right of priority under 35 U.S.C. §119 is hereby claimed.

In support of this claim, filed herewith is a certified copy of the original foreign application.

Respectfully submitted,

  
\_\_\_\_\_  
Gary R. Edwards  
Registration No. 31,824

CROWELL & MORING, LLP  
Intellectual Property Group  
P.O. Box 14300  
Washington, DC 20044-4300  
Telephone No.: (202) 624-2500  
Facsimile No.: (202) 628-8844  
GRE:kms  
(CAM 56207.522)

日 本 国 特 許 庁  
JAPAN PATENT OFFICE

11046 U.S. PTO  
10/083636  
02/27/02

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日

Date of Application:

2001年10月 2日

出 願 番 号

Application Number:

特願2001-306787

出 願 人

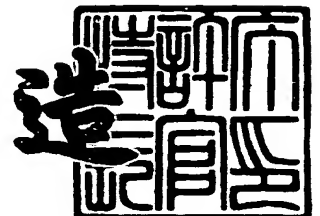
Applicant(s):

株式会社日立製作所

2001年11月30日

特 許 庁 長 官  
Commissioner,  
Japan Patent Office

及 川 耕 造



出証番号 出証特2001-3104862

【書類名】 特許願  
【整理番号】 PE28395  
【提出日】 平成13年10月 2日  
【あて先】 特許庁長官 殿  
【国際特許分類】 G10L 15/28  
G08G 1/0962

【発明者】

【住所又は居所】 茨城県日立市大みか町七丁目 1 番 1 号  
株式会社日立製作所 日立研究所内

【氏名】 葛貫 壮四郎

【発明者】

【住所又は居所】 茨城県日立市大みか町七丁目 1 番 1 号  
株式会社日立製作所 日立研究所内

【氏名】 大辻 信也

【発明者】

【住所又は居所】 茨城県日立市大みか町七丁目 1 番 1 号  
株式会社日立製作所 日立研究所内

【氏名】 森岡 道雄

【発明者】

【住所又は居所】 茨城県日立市大みか町七丁目 1 番 1 号  
株式会社日立製作所 日立研究所内

【氏名】 上脇 正

【発明者】

【住所又は居所】 茨城県日立市大みか町七丁目 1 番 1 号  
株式会社日立製作所 日立研究所内

【氏名】 奥出 真理子

【特許出願人】

【識別番号】 000005108

【氏名又は名称】 株式会社日立製作所

【代理人】

【識別番号】 100098017

【弁理士】

【氏名又は名称】 吉岡 宏嗣

【手数料の表示】

【予納台帳番号】 055181

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 音声入力システムおよび音声ポータルサーバおよび音声入力  
端末

【特許請求の範囲】

【請求項1】 音声入出力手段とWebブラウザと外部システムへのアクセス状況および検索結果を表示する表示手段とを備えた音声入力端末と、

前記音声入力端末からの音声を取り込みテキストとして認識する音声認識手段と、認識されたテキストをコマンドテキスト辞書と照合しコマンドテキストとオブジェクトテキストとに分離するコマンド変換手段と、分離されたコマンドテキストとオブジェクトテキストとに基づいて各種情報を提供するアプリケーションサービスプロバイダにアクセスしてサービスを受けこのサービスを前記音声入力端末に出力する対話制御手段とを備えた音声ポータルサーバと、

前記音声ポータルサーバから受信したコマンドテキストおよびオブジェクトテキストに基づいて情報を検索する情報検索手段を備え検索結果を前記音声ポータルサーバにサービスするアプリケーションサービスプロバイダと  
からなる音声入力システム。

【請求項2】 請求項1に記載の音声入力システムにおいて、

前記アプリケーションサービスプロバイダの情報検索手段が、入力されたオブジェクトテキストをn文字ごとに切り出し予め作成してあるn文字INDEXに基づいて情報検索する手段である  
ことを特徴とする音声入力システム。

【請求項3】 請求項1または2に記載の音声入力システムにおいて、

前記アプリケーションサービスプロバイダが、地図情報をサービスするナビゲーション情報アプリケーションサービスプロバイダである  
ことを特徴とする音声入力システム。

【請求項4】 請求項1ないし3のいずれか一項に記載の音声入力システムにおいて、

前記アプリケーションサービスプロバイダが、音楽情報をサービスする音楽情報アプリケーションサービスプロバイダである

ことを特徴とする音声入力システム。

【請求項5】 請求項1ないし4のいずれか一項に記載の音声入力システムにおいて、

前記アプリケーションサービスプロバイダが、TV番組、CS番組、CATV番組の少なくとも1つの情報をサービスする番組情報アプリケーションサービスプロバイダである

ことを特徴とする音声入力システム。

【請求項6】 請求項1ないし5のいずれか一項に記載の音声入力システムにおいて、

前記アプリケーションサービスプロバイダが、電話情報をサービスする電話情報アプリケーションサービスプロバイダである

ことを特徴とする音声入力システム。

【請求項7】 音声を入出力する手段と外部システムとのアクセス状況を表示する手段とを備えた音声入力端末と、各種情報を提供するアプリケーションサービスプロバイダと、入力された音声に基づいて前記音声入力端末とアプリケーションサービスプロバイダとの対話を制御する音声ポータルサーバとからなる音声入力システムにおいて、

前記音声ポータルサーバは、前記音声入力端末からの音声を取り込みテキストとして認識する音声認識手段と、認識されたテキストをコマンドテキスト辞書と照合しコマンドテキストとオブジェクトテキストとに分離するコマンド変換手段と、分離されたコマンドテキストとオブジェクトテキストとを前記アプリケーションサービスプロバイダに送り前記アプリケーションサービスプロバイダで検索された情報を音声入力端末に出力する対話制御手段とを備えた

ことを特徴とする音声ポータルサーバ。

【請求項8】 請求項7に記載の音声ポータルサーバにおいて、

前記音声認識手段が、連続音声認識手段と、単語音声認識手段と、入力された音声の特徴量をしきい値として前記2つの認識手段の認識結果の一方を選択して出力する認識総合評価手段とを備えた

ことを特徴とする音声ポータルサーバ。

【請求項 9】 請求項 8 に記載の音声ポータルサーバにおいて、  
前記音声の特徴量が、発話時間である  
ことを特徴とする音声ポータルサーバ。

【請求項 10】 請求項 8 に記載の音声ポータルサーバにおいて、  
前記音声の特徴量が、認識文字列長である  
ことを特徴とする音声ポータルサーバ。

【請求項 11】 音声ポータルサーバおよび各種情報を提供するアプリケーションサービスプロバイダにアクセスしサービスを受ける音声入力端末において

音声入出力手段と、Web ブラウザと、外部システムへのアクセス状況および  
検索結果を表示する表示手段とを備えた  
ことを特徴とする音声入力端末。

【請求項 12】 請求項 11 に記載の音声入力端末において、  
PDA、携帯電話、車載ナビゲーションのいずれかに組み込まれている  
ことを特徴する携帯音声入力端末。

【請求項 13】 請求項 11 に記載の音声入力端末において、  
家庭電話、TV、PC のいずれかに組み込まれている  
ことを特徴する家庭音声入力端末。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、音声入力システム、音声ポータルサーバ、音声入力端末に係り、特に、PDA(Personal Digital Assistants)、携帯電話、車載ナビゲーションなどの移動端末装置からまたは家庭電話、TV、PC などの家庭(固定)端末から音声でネットワークにアクセスし、地図情報、音楽情報、TV 番組情報、電話情報などを提供する情報提供サービスプロバイダから情報やサービスを受ける音声入力システムに関する。

【0002】

【従来技術】

特開平 1 1 - 1 4 3 4 9 3 号公報は、入力された音声を音声言語理解装置によりデータベース言語の中間言語に変換して単語を検索するシステムを記載している。

【 0 0 0 3 】

特開 2 0 0 0 - 5 7 4 9 0 号公報は、認識辞書を切り替えながら入力された音声の認識性能を向上させる方法を記載している。

【 0 0 0 4 】

特開 2 0 0 1 - 3 4 2 9 2 号公報は、ワードスポッティングという技術で辞書にある言葉を切り出し、要求キーワードを認識して話題を確定し、その話題用の認識辞書を用いて音声を認識させ、認識性能を向上させる方法を記載している。

【 0 0 0 5 】

【発明が解決しようとする課題】

特開平 1 1 - 1 4 3 4 9 3 号公報の技術は、識別誤りが最小になるように文章データをそれに対応する中間言語に変換する隠れマルコフモデルを学習する方法である。この方式では、統計的处理に基づく学習であるから、いろいろな分野に同時にサービスしようとする、分野ごとに学習する必要があり、処理に多くの時間がかかり、認識性能が下がる。また、長い文章と短い文章とが混在する実際の対話を考慮した音声入力システムにはなっていない。さらに、認識文字列の一部に誤りがあった場合についての配慮がない。

【 0 0 0 6 】

特開 2 0 0 0 - 5 7 4 9 0 号公報の技術は、認識結果に応じて対応辞書を切り替えて認識性能を向上させるナビゲーションの発明であるが、音声を連続入力できない。また、認識文字列の一部に誤りがあった場合についての配慮がない。

【 0 0 0 7 】

特開 2 0 0 1 - 3 4 2 9 2 号公報の技術は、認識結果に応じて話題を切り出し辞書を切り替えて認識性能を上げる発明である。しかし、上記 2 つの従来技術と同様に、認識文字列の一部に誤りがあった場合についての配慮がない。

【 0 0 0 8 】

本発明の目的は、PDA、携帯電話などの移動端末や家庭電話(Home T E L)、



TV、PCなどの固定端末から音声でネットにアクセスし、地図情報、音楽情報、番組情報、電話情報などを提供するプロバイダからサービスを受ける音声入力システムおよび音声ポータルサーバおよび音声入力端末を提供することである。

## 【0009】

## 【課題を解決するための手段】

本発明は、上記目的を達成するために、音声入出力手段とWebブラウザと外部システムへのアクセス状況および検索結果を表示する表示手段とを備えた音声入力端末と、音声入力端末からの音声を取り込みテキストとして認識する音声認識手段と、認識されたテキストをコマンドテキスト辞書と照合しコマンドテキストとオブジェクトテキストとに分離するコマンド変換手段と、分離されたコマンドテキストとオブジェクトテキストとに基づいて各種情報を提供するアプリケーションサービスプロバイダにアクセスしてサービスを受けこのサービスを音声入力端末に出力する対話制御手段とを備えた音声ポータルサーバと、音声ポータルサーバから受信したコマンドテキストおよびオブジェクトテキストに基づいて情報を検索する情報検索手段を備え検索結果を音声ポータルサーバにサービスするアプリケーションサービスプロバイダとからなる音声入力システムを提案する。

## 【0010】

アプリケーションサービスプロバイダの情報検索手段は、入力されたオブジェクトテキストをn文字ごとに切り出し予め作成してあるn文字INDEXに基づいて情報検索する手段とすることができる。

## 【0011】

アプリケーションサービスプロバイダは、地図情報をサービスするナビゲーション情報アプリケーションサービスプロバイダ、音楽情報をサービスする音楽情報アプリケーションサービスプロバイダ、TV番組、CS番組、CATV番組の少なくとも1つの情報をサービスする番組情報アプリケーションサービスプロバイダ、電話情報をサービスする電話情報アプリケーションサービスプロバイダなどとする。

## 【0012】

本発明においては、音声入力端末から入力された音声を音声ポータルサーバで

認識してコマンドとオブジェクトとのテキストに分離し、分離されたテキストに基づき、アプリケーションサービスプロバイダにある情報をあいまい検索し、オブジェクトテキストに一部認識誤りがあっても、意図する情報を音声入力端末に提供できる。

## 【0013】

本発明は、また、音声を入出力する手段と外部システムとのアクセス状況を表示する手段とを備えた音声入力端末と、各種情報を提供するアプリケーションサービスプロバイダと、入力された音声に基づいて音声入力端末とアプリケーションサービスプロバイダとの対話を制御する音声ポータルサーバとからなる音声入力システムにおいて、音声ポータルサーバは、音声入力端末からの音声を取り込みテキストとして認識する音声認識手段と、認識されたテキストをコマンドテキスト辞書と照合しコマンドテキストとオブジェクトテキストとに分離するコマンド変換手段と、分離されたコマンドテキストとオブジェクトテキストとをアプリケーションサービスプロバイダに送りアプリケーションサービスプロバイダで検索された情報を音声入力端末に出力する対話制御手段とを備えた音声ポータルサーバを提案する。

## 【0014】

音声認識手段は、連続音声認識手段と、単語音声認識手段と、入力された音声の特徴量をしきい値として2つの認識手段の認識結果の一方を選択して出力する認識総合評価手段とを備える。

## 【0015】

音声の特徴量は、発話時間、認識文字列長のいずれを採用してもよい。

## 【0016】

音声ポータルサーバの音声認識エンジンを長い文章に適した連続音声認識エンジンとコマンドのような短い文章に適した単語音声認識エンジンの2つで構成し、総合的に評価しているため、音声対話の認識性能が向上する。

## 【0017】

本発明は、さらに、音声ポータルサーバおよび各種情報を提供するアプリケーションサービスプロバイダにアクセスしサービスを受ける音声入力端末において

、音声入出力手段と、Webブラウザと、外部システムへのアクセス状況および検索結果を表示する表示手段とを備えている。

## 【0018】

音声入力端末は、PDA、携帯電話、車載ナビゲーションのいずれかに組み込まれている携帯音声入力端末と、家庭電話、TV、PCのいずれかに組み込まれている家庭音声入力端末とに分類できる。

## 【0019】

アプリケーションサービスプロバイダ(ASP)として、ナビゲーション情報ASP、音楽情報ASP、番組情報ASP、電話情報ASPを設けたので、PDA、MobileTEL、MobileCarPCなどの移動する音声入力端末や、家庭電話、TV、PCなどの家庭の音声入力端末にも、それぞれの必要に応じて、最適な情報をサービスできる。

## 【0020】

## 【発明の実施の形態】

次に、図1～図28を参照して、本発明による音声入力システム、音声ポータルサーバ、音声入力端末の実施形態を説明する。

## 【0021】

図1は、本発明による音声入力システムの実施形態の全体構成を示すブロック図である。

## 【0022】

本実施形態では、音声入力端末装置としては、移動端末10と家庭(固定)端末30とを使用できる。移動端末10としては、PDA10a、携帯電話10b、車載端末10cなどがある。家庭(固定)端末30としては、家庭内での利用を想定した固定電話30a、情報家電としてのテレビシステムTV30b、パーソナルコンピュータPC30cなどがある。移動端末10a～10cは、無線基地局20を介してインターネット網40に接続され、家庭端末30a～30cは、インターネット網40に直接接続されている。

## 【0023】

音声対話全体を制御する音声ポータル50および各種アプリケーションサービ

スプロバイダ(ASP)60も、インターネット網40に接続されている。

【0024】

ASPとしては、地図情報をサービスするナビゲーション情報ASP60a、音楽情報ASP60b、TV番組情報ASP60c、電話情報ASP60dなどがある。

【0025】

音声入力端末10a~10c、30a~30cのいずれかから、音声ポータルサーバ50に接続すると、音声のガイダンスやメニュー表示が音声入力端末に出力され、対応する音声を入力すると、音声はインターネット網40を介して音声ポータルサーバ50に転送される。

【0026】

音声ポータルサーバ50は、音声を認識し、音声の内容をコマンドや検索対象物であるオブジェクトとしてコマンド変換処理し、コマンドの内容に対応したASP60に転送する。

【0027】

ASP60では、対応するデータベースを検索し、音声ポータルサーバ50を経由して、音声を入力してきた音声入力端末に検索結果を出力する。

【0028】

以上に説明したように、音声入力システムは、キーボード(KB)を利用しにくい環境にある移動端末やキーボード操作に慣れていない家庭での端末に主に利用され、入力を容易にする。

【0029】

本実施形態においては、音声入力システムの全体構成として、サーバ群をインターネットに接続してある。

【0030】

サーバ群をイントラネットやホームネットに接続しても、これらのネットワーク内では、効果は変わらない。各種ASP群を近くに置き、これらのASP群では意図するデータをサービスできないときのみインターネットのサーバ群に接続するいわゆるキャッシュサーバとして設置してもよい。

【0031】

なお、ASPサーバ群には、図1に示した以外の情報サービス、例えば、株価情報、取引先情報、顧客情報、商品情報などが存在してもよい。

【0032】

音声ポータルサーバ50は、個人情報进行管理し、個人の特性に応じたサービスをすることも可能である。

【0033】

図2～図4は、移動端末の構成を示す図であり、図5～図7は、家庭(固定)端末の構成を示す図である。各端末の基幹部分は、ほぼ同じように構成されている。

【0034】

図2は、本発明による音声入力端末であるPDAの実施形態の構成を示すブロック図である。PDA10aは、無線基地局20と通信するアンテナ10a1と、無線通信するための通信手段10a2とを含んでいる。通信手段10a2は、Voice Over IP(VoIP)技術などにより、音声とデータとを同時に送受信できる。処理装置およびWebブラウザ10a3は、各構成部および周辺装置と接続され、端末全体を制御する。周辺装置としては、音声入力用のマイクMIC10a4、タッチパネルとして構成する座標入力装置(タブレット)TB10a5、液晶ディスプレイLCD10a6、スピーカSP10a7がある。

【0035】

PDA10aは、移動端末にとって重要な位置検出手段10a8を備えており、GPS(Global Positioning System)10a9に接続されている。

【0036】

PDA10aの操作は、タッチパネルと音声でなされる。その処理結果はディスプレイに表示され、次の操作の待ち状態となる。

【0037】

図3は、本発明による音声入力端末であるMobileTELの構成を示すブロック図である。構成要素は、図2のPDA10aと同じである。しかし、一般に、価格を抑えるため、液晶ディスプレイLCD10b6のサイズやカラー表示性能が

異なる。一方、携帯電話用各種応用ソフトが追加されている。

【0038】

図4は、本発明による音声入力端末であるMobileCar P Cの構成を示すブロック図である。構成要素は、基本的には、図2のPDA 10 aと同じである。しかし、車載用に適した液晶ディスプレイLCD 10 c 6や車載用応用ソフトが、図2のPDA 10 aとは異なる。また、図示していないが、車載の各種センサと接続され、車両に関する情報を表示してもよい。

【0039】

図5は、本発明による音声入力端末である家庭電話の構成を示すブロック図である。図2のPDA 10 aとの違いは、無線基地局20との通信用アンテナ10 a 1、位置検出手段10 a 8、GPS 10 a 9が無いことである。

【0040】

図6は、本発明による音声入力端末であるTVの構成を示すブロック図である。図5の固定電話家庭電話30 aとは、テレビ装置TV 30 b 10、TV制御手段30 b 8、カメラCM 30 c 9が異なっている。TV制御手段30 b 8は、TV番組を予約したり、チャンネルを設定したりするための手段であり、一般に、セットトップボックスと呼ばれる。

【0041】

カメラCM 30 c 9は、通信相手に会話用映像を送ったり、室内を画像で監視したりすることに用いる。

【0042】

図7は、本発明による音声入力端末であるP Cの構成を示すブロック図である。図6のテレビシステムTV 30 bと異なるのは、TV制御手段が無いことである。P Cの操作は、タッチパネルまたは音声でなされる。図示しないが、キーボードを接続し、P Cを操作してもよい。

【0043】

なお、図6および図7に示したカメラCM 30 c 9は、図2～図5の音声入力端末に装備してもよい。

【0044】

図 8 は、本発明による音声ポータルサーバ 5 0 の実施形態の構成を示すブロック図である。本発明の特徴的な部分である音声ポータルサーバ 5 0 は、インターネット 4 0 と通信する通信手段 5 0 1 と、音声ポータルサーバ 5 0 全体を処理する処理装置 5 0 2 と、音声データ Vin を取り込み認識辞書 5 0 4 を用いてテキストデータ Vtext 1 に認識出力する音声認識手段 5 0 3 と、認識された音声 Vtext 1 をコマンドテキスト辞書 5 0 6 を用いてコマンドとオブジェクト Vtext 2 に変換するコマンド変換手段 5 0 5 と、音声入力端末や各種情報 A S P との対話を制御する対話制御手段 5 0 7 と、対話制御手段の音声テキスト Vtext 3 により音声を合成する音声合成手段 5 0 8 と、Web ブラウザ 5 0 9 とからなる。

## 【 0 0 4 5 】

図 9 は、本発明による音声ポータルサーバ 5 0 の音声認識手段 5 0 3 の構成を示すブロック図である。本実施形態では、音声認識手段が 2 つの認識エンジンを備えていることが特徴である。すなわち、音声認識手段 5 0 3 は、比較的長い音声を認識させる連続音声認識エンジン 5 0 3 a とコマンドのような比較的短い音声を認識する単語音声認識エンジン 5 0 3 b で構成される。

## 【 0 0 4 6 】

連続音声認識エンジン 5 0 3 a は、連続音声認識辞書 5 0 4 a を用いて音声を認識し、単語音声認識辞書エンジン 5 0 3 b は、単語音声認識辞書 5 0 4 b を用いて音声を認識する。

## 【 0 0 4 7 】

各認識エンジンの認識結果は、認識総合的評価手段 5 0 3 c で総合的に評価される。一般に、連続音声認識エンジンは、単語間の遷移確率モデルを用いて認識する方式を用いるため、コマンドのような短い単語が入力されると、前後の単語知識を利用できないため、誤認識が多くなる。

## 【 0 0 4 8 】

そこで、認識総合評価手段で、どちらの認識エンジンの出力が正しいかを総合的に判断することが必要になる。

## 【 0 0 4 9 】

認識総合的評価手段 5 0 3 c の動作を図 1 0 ～図 1 3 を用いて具体例で説明す

る。

#### 【0050】

図10～図11は、2つの認識エンジンの認識結果を音声の発話時間を用いて切り替える例を示している。認識総合的評価手段503cは、発話時間評価において、音声データVinをしきい値と比較し、短ければS側に、長ければL側に切り替える。

#### 【0051】

図10は、本発明による音声ポータルサーバ50の音声認識手段の動作を説明する図であり、発話時間が比較的短い“はい。”という音声が入力されたときの状態を説明する図である。この場合は、S側に切り替えられて、認識総合評価手段の出力Vtext1は、“はい。”の文字列が出力される。ここで、しきい値は、単語音声認識辞書の最大発話時間を選んでおく。

#### 【0052】

図11は、本発明による音声ポータルサーバ50の音声認識手段の動作を説明する図であり、発話時間が比較的長い“日立神田に目的地を設定する。”という音声発話データが入力されたときの状態を示す図である。この場合は、L側に切り替えられて、認識総合評価手段の出力Vtext1は、“日立神田に目的地を設定する。”の文字列が出力される。

#### 【0053】

図12～図13は、発話時間を評価するのではなく、認識エンジンの結果の文字列長をしきい値と評価する例である。

#### 【0054】

図12は、本発明による音声ポータルサーバ50の音声認識手段の動作を説明する図である。発話時間が比較的短い“はい。”という音声が入力されたときの状態を説明する図である。この場合は、S側に切り替えられて、認識総合評価手段の出力Vtext1は、“はい。”の文字列が出力される。

#### 【0055】

図13は、本発明による音声ポータルサーバ50の音声認識手段の動作を説明する図である。発話時間が比較的長い“日立神田に目的地を設定する。”という



音声発話データが入力されたときの状態を示す図である。この場合は、L側に切り替えられて、認識総合評価手段の出力Vtext1は、“日立神田に目的地を設定する。”の文字列が出力される。

## 【0056】

なお、認識エンジンでは、余りにも辞書とかけ離れた音声が入力されると“認識不能を示す文字列＝“？””を出力するようにしている。この場合においても、しきい値を適当な値(例：コマンド文字列の最大長)を選んでおくことにより、最適な文字列が出力され、全体の認識性能が改善される。

## 【0057】

以上の2つの方式で説明したように、音声発話データが“はい。”というコマンドであっても、連続音声認識エンジンのみの場合、“肺。”という文字が出力されてしまう問題を解決できる。

## 【0058】

図14は、本発明による音声ポータルサーバ50のコマンド変換手段505の構成を示すブロック図である。音声認識手段503の文字列Vtext1がコマンド変換手段に入力されると、コマンド文字列検索505aにおいて、コマンドテキスト辞書506を用いて、コマンド文字列が含まれるかどうかを処理する。

## 【0059】

図15は、本発明による音声コマンドテキスト辞書の例を示す図である。コマンドテキスト辞書506では、コマンドID、コマンド名1～コマンド名5まであり、同一コマンドIDでは、どれが入力されてもよいようになっている。例えば、“目的地設定。”と“行く。”の文字列は同一コマンドID=D01として判断される。

## 【0060】

図15のコマンドテキスト文字は、各情報ASPにあわせたコマンド(NO1～NO8)、音声対話用のコマンド(NO9～NO10)、画面操作用のコマンド(NO11～NO22)に大別できる。

## 【0061】

コマンド文字列検索は、完全一致を前提としているが、図23～図24で後述す

るように、一部認識誤りを含んでいても検索できるようにしてもよい。

【0062】

コマンド文字列検索505a処理後は、コマンド以外の文字列であるオブジェクトを抽出するオブジェクト抽出505bを実行する。この処理は、各情報ASPに送出するコマンドと検索対象文字列であるオブジェクトとを抽出する処理である。

【0063】

図16は、本発明による音声ポータルサーバ50のコマンド変換手段505の動作を説明する図である。音声認識手段の結果Vtext1が、“日立神田に目的地を設定する。”であるとする、コマンド文字列検索505aでは、コマンドテキスト辞書506を参照してコマンド文字列が“目的地を設定する。”と判断し、コマンドID=D01と理解する。

【0064】

次に、オブジェクト抽出505bでは、コマンド文字列以外のがオブジェクトと判断し、オブジェクト“日立神田に。”が抽出される。したがって、オブジェクト抽出手段505bの結果Vtext2は、“コマンドID=D01、オブジェクト=日立神田に。”として出力される。

【0065】

なお、オブジェクト抽出手段505bでは、処理を簡単にするため、コマンド文字列以外の全ての文字列をオブジェクトにしたが、形態素解析を実行し、“日立神田に。”の“に”を省いて抽出するようにしてもよい。

【0066】

図17は、本発明による音声ポータルサーバ50の対話制御手段507の構成を示すブロック図である。対話制御手段507は、全体を制御する対話処理手段507aと、音声対話の基本対話ルール507bと、音声入力端末とのインタフェースである端末データ制御手段507cと、各情報ASPとのインタフェースであるASP制御手段507dと、音声合成制御手段507eとで構成される。

【0067】

基本対話ルール507bは、各情報ASPで共通に使用されるルールを格納す

るものであり、各情報ASPの特有の対話ルールは、各情報ASPからダウンロードされる。

## 【0068】

まず、音声認識手段505からの出力であるVtext2が入力されると、対話処理手段507aで、コマンドIDを判断し、どの情報ASPに対応するかを判断し、対応する情報ASPにコマンドIDおよびオブジェクトをASPDataOutとして送信する。

## 【0069】

次に、対応した情報ASPからの検索結果が、ASPDataInとしてASP制御手段507bに入力されたら、端末制御507cは、その検索を要求した音声入力端末に対して、TdataOutを出力する。音声入力端末では、その検索結果のデータが表示される。

## 【0070】

さらに、文字列により音声合成出力する場合は、音声合成制御手段507eから、Vtext3として音声列を出力し、音声合成手段508で合成した音声Voutを音声入力端末に送り、スピーカから音を出す。

## 【0071】

なお、音声入力端末からの音ではないデータ入力がある場合は、TdataInとして受け付ける。

## 【0072】

また、対話制御手段507は、音声対話用VoiceXMLブラウザで構成してもよい。

## 【0073】

次に、図18～図21により、各情報ASPの詳細な構成を説明する。

## 【0074】

図18は、本発明によるナビゲーション情報ASPの構成を示すブロック図である。ナビゲーション情報ASPは、地図情報や経路探索情報をサービスするプロバイダであり、インターネットとのインタフェース60a100と、あいまい検索手段60a200と、経路探索手段60a500と、対話ルール処理手段6

0a700とからなる。各手段は、各々辞書を参照して要求を処理する。

【0075】

あいまい検索手段60a200は、行き先情報であるランドマーク情報のデータベースランドマークDB60a300と、あいまい検索するためのランドマークINDEX60a400辞書とを参照する。詳細な動作については、後述する。

【0076】

経路探索手段60a500は、地図データであるMapDB60a600を参照して現在地から、目的地までの経路を探索する。この経路探索は、一般になされている経路探索処理であるので、詳細な説明を省略する。

【0077】

対話ルール処理手段60a700は、各情報ASPごとに特有の対話ルールを処理する手段であり、対話ルール60b800は、音声ポータルサーバ50の基本対話以外のルールとして利用される。

【0078】

図19は、本発明による音楽情報ASPの構成を示すブロック図である。図18のASPと比較すると、音楽情報ASPには、経路探索手段に相当するものが無く、コンテンツが、音楽DB60b300、音楽INDEX60b400、音楽用対話ルール60b800となっている。

【0079】

図20は、本発明によるTV番組情報ASPの構成を示すブロック図である。図19のASPとは、コンテンツの内容が異なるだけである。コンテンツとしては、TV番組DB60c300、番組INDEX60c400、番組用対話ルール60c800がある。

【0080】

なお、本明細書で、TV番組とは、TV番組、CS番組、CATV番組の少なくとも1つの情報を意味する。

【0081】

図21は、本発明による電話情報ASPの構成を示すブロック図である。

図19のASPとは、コンテンツが異なり、電話DB60d300、電話INDEX60d400、電話用対話ルール60d800を備えている。

【0082】

図22は、本発明による音声操作メニュー画面の一例を示す図である。音声メニューには、各情報ASPに応じて音声メニューアイコンが用意されている。ナビゲーション情報ASPに関するものには、“目的地設定。”，“目的地検索。”，“経由地設定。”，“現在地表示。”がある。

【0083】

音楽情報ASPに関するものには、“音楽検索。”がある。番組情報ASPに関するものには、“番組検索。”、“番組予約。”がある。電話情報ASPに関するものには、“電話検索。”がある。

【0084】

本発明では、項目を音声メニューから入力する方法と、対象物も含めてすべて音声で入力する方法も許している。例えば、目的地検索の例では、メニューを押さずに、“日立神田に目的地を設定する。”と直接音声入力してもよい。

【0085】

図23は、本発明による各種情報ASPのあいまい検索手段60a200の構成を示すブロック図である。他のあいまい検索手段60b200、60c200、60d200の構成も、あいまい検索手段60a200と同様である。

【0086】

図23において、あいまい検索手段60a200は、検索エンジン60a100と、2文字INDEX生成60a220とからなる。検索エンジン60a100および2文字INDEX生成60a220は、ランドマークDBa300およびランドマークINDEX60a400を参照しながら検索処理する。

【0087】

ランドマークDBは、数百万件の膨大なデータを保有しているので、ランドマークINDEXは、2文字INDEX生成処理により、予め作成しておく必要がある。本発明においては、この2文字INDEX生成処理により、高速でかつあいまい検索がなされるのが特徴である。ここで、あいまい検索とは、検索語句の

意味があいまいではなく、文字列の一部の誤り(一部文字列追加、一部文字列欠損、文字列順不同、一部文字列誤り)があっても、入力された語句を検索できることを意味している。

【0088】

図24は、本発明による各種情報ASPのあいまい検索の手順の一例を示す図である。例として、“日立神田に。”に目的地設定をする場合を示す。

【0089】

まず、“日立神田に。”の検索キーワード60a200keyが入力されると、2文字ごとに切出処理60a211をする。

【0090】

次に、2文字ごとにランドマークINDEX検索60a212をする。

【0091】

ランドマークINDEX検索60a212からランドマークDBを検索し、ヒットするDBのレコードを抽出する。

【0092】

抽出したレコードのうち、ヒット文字数の多いものの順にソーティングし出力処理60a214し、検索結果としてリスト60a200resを出力する。

【0093】

以上のように、2文字ごとのINDEX検索するので、高速検索とあいまい検索の両方の特徴をもつ。

【0094】

ここで、“日立神田に。”の“に。”が検索オブジェクトにあってもランドマークDBに該当するものがなければ、無視される。一方、“神田日立に。”と入力されても、関係するものがヒットする。

【0095】

このため、思いついた地名やランドマークを入力できる効果がある。

【0096】

なお、複数の検索結果を音声入力端末に出力し、どの検索対象物を選択するかを音声指示し、選択することもできる。

## 【0097】

本実施形態では、図23～図24に図示したように、2文字INDEX生成処理により、検索する例を示したが、3文字INDEXや4文字INDEX生成でもよい。

## 【0098】

特に、数字やアルファベットが多い情報をもつ場合、3文字INDEXや4文字INDEX生成の方が、2文字INDEX処理よりも、不要な検索出力が少なくなるという効果がある。

## 【0099】

図25～図28を用いて、音声入力端末、音声ポータルサーバ50、情報ASP60間の具体的な通信手順を説明する。

## 【0100】

図25は、本発明による音声入力端末と音声ポータルサーバとナビゲーション情報ASPとの間の通信手順を示す図である。ここでは、音声入力端末MobilePC10c、音声ポータルサーバ50、ナビゲーション情報ASP60a間の通信手順を示すが、他の情報ASP間との通信手順もほぼ同様である。

## 【0101】

まず、音声入力端末MobilePC10cから音声ポータルサーバ50に接続要求が出されると、音声ポータルサーバ50から、音声により“ご用件を入力ください。”と音声入力端末MobilePC10cに音声出力する。同時に、図21の音声メニューも表示される。

## 【0102】

次に、音声入力端末MobilePC10cでは、音声で“日立神田に目的地を設定する。”と直接音声入力する。

## 【0103】

音声ポータルサーバ50では、これを認識し、対応して“日立神田に目的地を設定しますか？”と音声応答する。

## 【0104】

ここで、“はい。”または“いいえ。”のコマンド用の音声を入力する。

## 【0105】

ここでは、“はい。”が入力されると、音声ポータルサーバ50は、“検索中です。”の音声応答を音声入力端末MobilePC10cに返すとともに、ナビゲーション情報ASP60aに対し、コマンドID“D01。”、オブジェクト“日立神田に。”のデータを送信し、検索結果を返してもらう。ここでは、検索結果数(2件)と内容(XXX, YYY)を返す。

## 【0106】

そこで、音声ポータルサーバ50は、この検索結果から“結果が2件あります。何番にしますか？”と音声応答する。同時に、音声入力端末MobilePC10cのディスプレイには、検索結果の内容表示がされる。

## 【0107】

次に、音声で“1番。”と音声指示すると、音声ポータルサーバ50は音声を認識し、対応する音声“目的地を1番に設定します。”を出力する。

## 【0108】

さらに、音声入力端末MobilePC10cに現在位置を要求し、現在位置情報を得て、この情報を元に、ナビゲーション情報ASP60aに対して、経路探索コマンドとそのパラメータとを送信する。

## 【0109】

音声ポータルサーバ50は、ナビゲーション情報ASPから、探索結果である経路情報および地図情報を得て、音声入力端末MobilePC10cに出力するとともに、音声“安全運転でどうぞ。”と音声応答する。

## 【0110】

ここで、上記通信手順で、“いいえ。”が入力されると、図示していないが、“ご用件を入力ください。”に再度戻るようになっている。

## 【0111】

複数ある検索結果を選択する場合、音声で選択する例を示したが、タッチパネルを備えているので、タッチパネルで選択してもよい。この場合、検索結果の内容とタッチパネルの座標との対応は、予め決めておかなければならない。

## 【0112】



図 2 5 において、音声入力端末として Mobile P C 1 0 c を用いたが、P D A 1 0 a、Mobile T E L 1 0 b とナビゲーション情報 A S P 間でやり取りしてもよい。この場合、人間ナビゲーションシステムになる。

すなわち、自分の音声入力端末の現在位置がわかるので、現在位置情報を表示したり、行きたいランドマークを検索したりできる。

【 0 1 1 3 】

図 2 6 は、本発明による音声入力端末と音声ポータルサーバと音楽情報 A S P との間の通信手順を示す図である。音声入力端末 Mobile P C 1 0 c が、音声ポータルサーバ 5 0 を介して音楽情報 A S P 6 0 b から音楽のコンテンツサービスを受ける通信手順である。

【 0 1 1 4 】

まず、音声入力端末 Mobile P C 1 0 c から音声ポータルサーバ 5 0 に接続要求が出されると、音声ポータルサーバ 5 0 から、音声により“ご用件を入力ください。”と音声入力端末 Mobile P C 1 0 c に音声出力する。同時に、図 2 1 の音声メニューも表示される。

【 0 1 1 5 】

次に、音声入力端末 Mobile P C 1 0 c では、音声で“Mariah Careyの曲を聴きたい。”と直接音声入力する。

【 0 1 1 6 】

音声ポータルサーバ 5 0 では、その音声を認識し、対応して“Mariah Careyの曲を検索しますか？。”と音声応答がある。

【 0 1 1 7 】

ここで、“はい。”または“いいえ。”のコマンド用の音声を入力する。

【 0 1 1 8 】

ここでは、“はい。”が入力されると、音声ポータルサーバ 5 0 は、“検索中です。”の音声応答を音声入力端末 Mobile P C 1 0 c に返すとともに、音楽情報 A S P 6 0 b に対し、コマンド I D “M01.”、オブジェクト“Mariah Careyの。”のデータを送信し、検索結果を返してもらう。

【 0 1 1 9 】

ここでは、検索結果数(3件)と内容(XXX, YYY, ZZZ)を返す。

【0120】

そこで、音声ポータルサーバ50は、この検索結果から“結果が3件あります。何番にしますか?”と音声応答する。同時に、音声入力端末MobilePC10cのディスプレイには、検索結果の内容表示がされている。

【0121】

次に、音声で“3番。”と音声指示すると、音声ポータルサーバ50はこれを認識し、対応する音声“音楽を3番に設定します。”と出力する。同時に、3番の音楽のダウンロードを音楽情報ASPに指示する。

【0122】

これにより、“お楽しみください。”の音声応答と対応した音楽コンテンツが音声入力端末MobilePC10cにダウンロードされる。

【0123】

もし、検索結果が1件しか無い時のとき、OKかどうかの音声応答を待って、ダウンロードが開始する。

【0124】

図27は、本発明による音声入力端末と音声ポータルサーバとTV番組情報ASPとの間の通信手順を示す図である。音声入力端末TV30b、PC30cが、音声ポータルサーバ50を介して番組情報ASP60cからTV番組のコンテンツサービスを受ける通信手順である。

【0125】

まず、音声入力端末から音声ポータルサーバ50に接続要求が出されると、音声ポータルサーバ50から、音声により“ご用件を入力ください。”と音声入力端末に音声出力する。同時に、図21の音声メニューも表示される。

【0126】

次に、音声入力端末では、音声で“天気予報の番組を見たい。”と直接音声入力する。

【0127】

音声ポータルサーバ50では、これを認識し、対応して“天気予報の番組を検

索しますか?。”と音声応答がある。

【0128】

ここで、“はい。”または“いいえ。”のコマンド用の音声を入力する。

【0129】

ここでは、“はい。”が入力されると、音声ポータルサーバ50は、“検索中です。”という音声応答を音声入力端末に返すとともに、番組情報ASP60cに対し、コマンドID“T01”、オブジェクト“天気予報の。”のデータを送信し、検索結果を返してもらう。

【0130】

ここでは、検索結果数(2件)と内容(XXX, YYY)を返す。

【0131】

そこで、音声ポータルサーバ50は、この検索結果から“結果が2件あります。何番にしますか?”と音声応答する。同時に、音声入力端末のディスプレイには、検索結果の内容表示がされる。

【0132】

次に、音声で“1番。”と音声指示すると、音声ポータルサーバ50はこれを認識し、対応する音声“番組を1番に設定します。”を出力する。

【0133】

その結果、TV番組に対応したチャンネルが設定されて、天気予報サービスを見ることができる。

【0134】

もし、複数の番組がなければ、OKかどうかの音声応答を待つてチャンネルを設定する。

【0135】

ここで、もし、天気予報サービスが放送されていないときは、チャンネルを予約できる。この場合は、音声ポータルサーバ50から、予約するかどうかの案内があり、それに答えると、予約が終了する。

【0136】

なお、毎週見ているTV番組については、毎週予約などもできる。

## 【0137】

図28は、本発明による音声入力端末と音声ポータルサーバと電話情報ASPとの間の通信手順を示す図である。音声入力端末家庭電話30a, MobileTEL10bが、音声ポータルサーバ50を介して、電話情報ASP60dから電話情報のコンテンツサービスを受ける通信手順である。

## 【0138】

まず、音声入力端末から音声ポータルサーバ50に接続要求が出されると、音声ポータルサーバ50から、音声により“ご用件を入力ください。”と音声入力端末に音声出力する。同時に、図21の音声メニューも表示される。

## 【0139】

次に、音声入力端末では、音声で“日立太郎さんに電話をしたい。”と直接音声入力する。

## 【0140】

音声ポータルサーバ50では、これを認識し、対応して“日立太郎さん電話をかけますか？”と音声応答する。

## 【0141】

ここで、“はい。”または“いいえ。”のコマンド用の音声を入力する。

## 【0142】

ここでは、“はい。”が入力されると、音声ポータルサーバ50は、“検索中です。”の音声応答を音声入力端末に返すとともに、電話情報ASP60dに対し、コマンドID“P01.”、オブジェクト“日立太郎さんに。”のデータを送信し、検索結果を返してもらう。

## 【0143】

ここでは、検索結果数(2件)と内容(XXX, YYY)を返す。

## 【0144】

そこで、音声ポータルサーバ50は、この検索結果から“結果が2件あります。何番にしますか？”と音声応答する。同時に、音声入力端末のディスプレイには、検索結果のコンテンツが表示されている。

## 【0145】

次に、音声で“1番。”と音声指示すると、音声ポータルサーバ50はこれを認識し、対応する音声“電話を1番にかけます。”と出力する。

【0146】

具体的には、日立太郎さんの電話に電話をかけて、相手がでたら対話できる。もし、検索結果が一人だけならば、OKかどうかの音声応答を待って電話をかける。

【0147】

【発明の効果】

本発明によれば、音声入力端末から入力された音声を音声ポータルサーバで認識してコマンドとオブジェクトとのテキストに分離し、分離されたテキストに基づき、アプリケーションサービスプロバイダにある情報をあいまい検索し、オブジェクトテキストに一部認識誤りがあっても、意図する情報を音声入力端末に提供できる。

【0148】

また、音声ポータルサーバの音声認識エンジンを長い文章に適した連続音声認識エンジンとコマンドのような短い文章に適した単語音声認識エンジンの2つで構成し、総合的に評価しているため、音声対話の認識性能が向上する。

【0149】

さらに、アプリケーションサービスプロバイダ(ASP)として、ナビゲーション情報ASP、音楽情報ASP、番組情報ASP、電話情報ASPを設けたので、PDA、MobileTEL、MobileCarPCなどの移動する音声入力端末や、家庭電話、TV、PCなどの家庭の音声入力端末にも、それぞれの必要に応じて、最適な情報をサービスできる。

【図面の簡単な説明】

【図1】

本発明による音声入力システムの実施形態の全体構成を示すブロック図である。

【図2】

本発明による音声入力端末であるPDAの実施形態の構成を示すブロック図で

ある。

【図 3】

本発明による音声入力端末である Mobile T E L の構成を示すブロック図である。

【図 4】

本発明による音声入力端末である Mobile Car P C の構成を示すブロック図である。

【図 5】

本発明による音声入力端末である家庭電話の構成を示すブロック図である。

【図 6】

本発明による音声入力端末である T V の構成を示すブロック図である。

【図 7】

本発明による音声入力端末である P C の構成を示すブロック図である。

【図 8】

本発明による音声ポータルサーバの実施形態の構成を示すブロック図である。

【図 9】

本発明による音声ポータルサーバの音声認識手段の構成を示すブロック図である。

【図 1 0】

本発明による音声ポータルサーバの音声認識手段の動作を説明する図である。

【図 1 1】

本発明による音声ポータルサーバの音声認識手段の動作を説明する図である。

【図 1 2】

本発明による音声ポータルサーバの音声認識手段の動作を説明する図である。

【図 1 3】

本発明による音声ポータルサーバの音声認識手段の動作を説明する図である。

【図 1 4】

本発明による音声ポータルサーバのコマンド変換手段の構成を示すブロック図である。

【図 1 5】

本発明による音声コマンドテキスト辞書の例を示す図である。

【図 1 6】

本発明による音声ポータルサーバのコマンド変換手段の動作を説明する図である。

【図 1 7】

本発明による音声ポータルサーバの対話制御手段の構成を示すブロック図である。

【図 1 8】

本発明によるナビゲーション情報 A S P の構成を示すブロック図である。

【図 1 9】

本発明による音楽情報 A S P の構成を示すブロック図である。

【図 2 0】

本発明による T V 番組情報 A S P の構成を示すブロック図である。

【図 2 1】

本発明による電話情報 A S P の構成を示すブロック図である。

【図 2 2】

本発明による音声操作メニュー画面の一例を示す図である。

【図 2 3】

本発明による各種情報 A S P のあいまい検索手段の構成を示すブロック図である。

【図 2 4】

本発明による各種情報 A S P のあいまい検索の手順の一例を示す図である。

【図 2 5】

本発明による音声入力端末と音声ポータルサーバとナビゲーション情報 A S P との間の通信手順を示す図である。

【図 2 6】

本発明による音声入力端末と音声ポータルサーバと音楽情報 A S P との間の通信手順を示す図である。

【図 27】

本発明による音声入力端末と音声ポータルサーバとTV番組情報ASPとの間の通信手順を示す図である。

【図 28】

本発明による音声入力端末と音声ポータルサーバと電話情報ASPとの間の通信手順を示す図である。

【符号の説明】

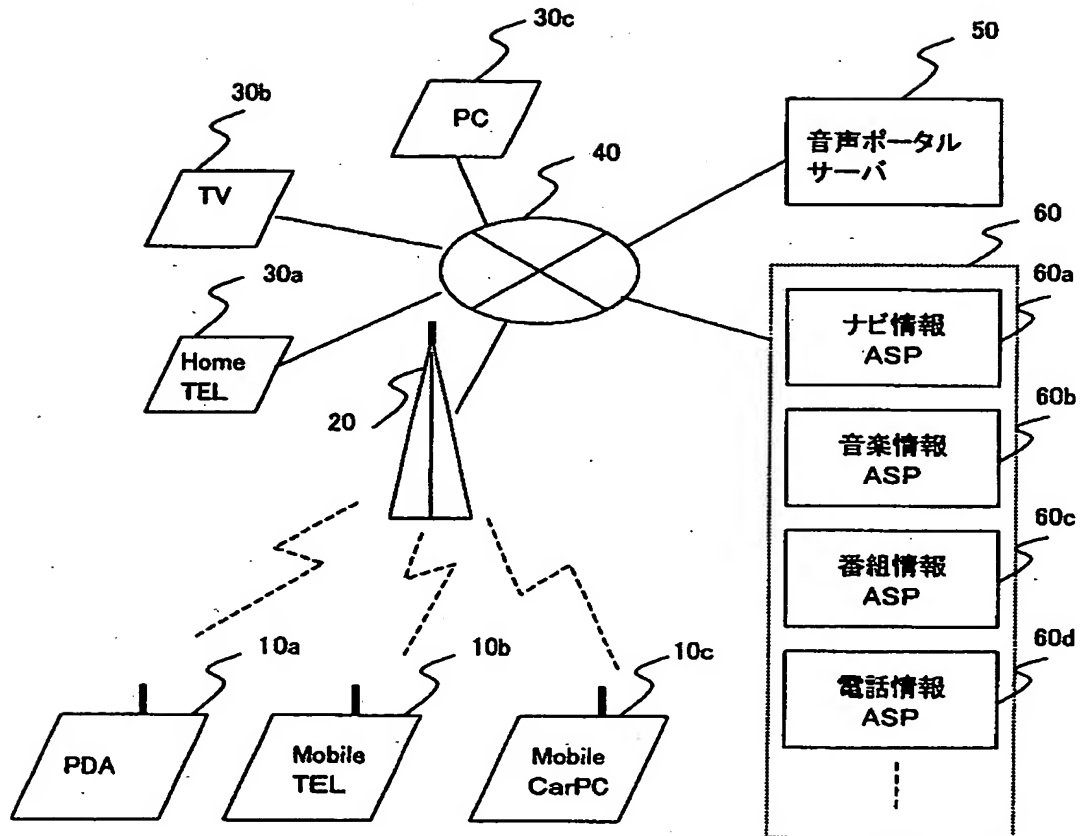
- 10 携帯端末
  - 10a PDA
  - 10b MobileTEL
  - 10c MobileCarPC
- 20 無線基地局
- 30 家庭端末
  - 30a 家庭電話HomeTEL
  - 30b TV
  - 30c PC
- 40 インターネット網
- 50 音声ポータルサーバ
  - 501 通信手段
  - 502 処理装置
  - 503 音声認識手段
    - 503a 連続音声認識
    - 503b 単語音声認識
    - 503c 認識総合評価
  - 504 認識辞書
    - 504a 連続音声認識辞書
    - 504b 単語音声認識辞書
  - 505 コマンド変換手段
    - 505a コマンド文字列検索



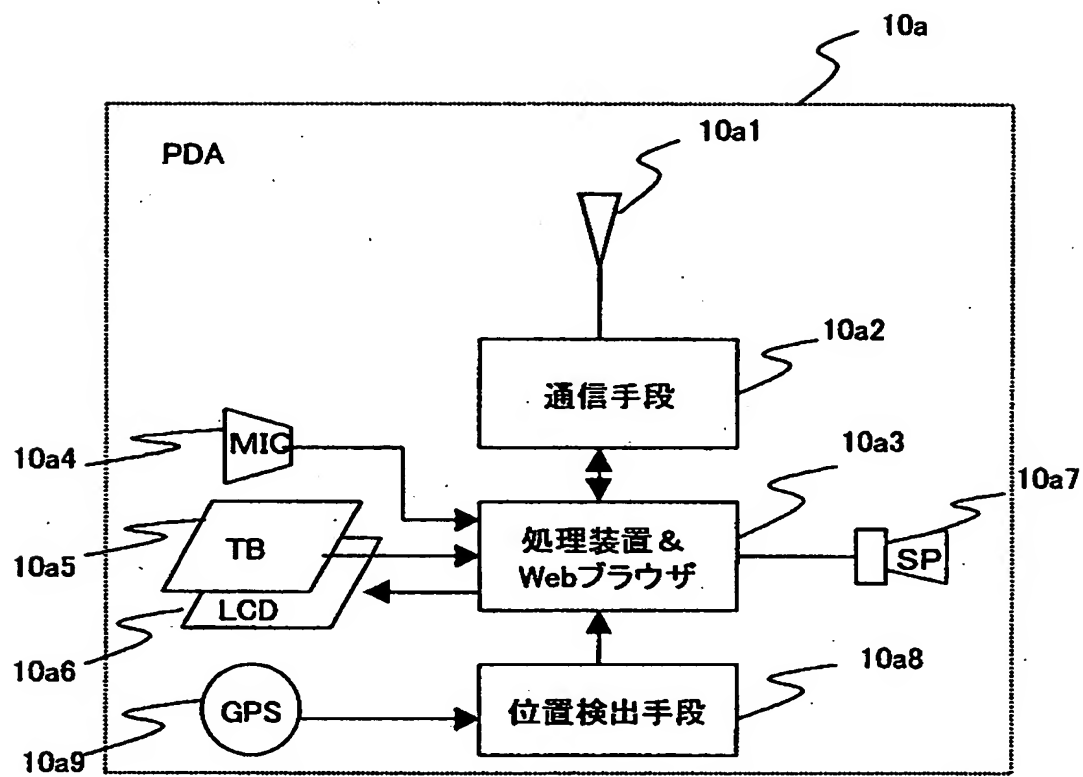
- 505b オブジェクト抽出
- 506 コマンドテキスト辞書
- 507 対話制御手段
  - 507a 対話処理手段
  - 507b 基本対話ルール
  - 507c 端末データ制御
  - 507d ASP制御手段
  - 507e 音声合成制御
- 508 音声合成手段
- 509 Webブラウザ
- 60 情報ASP
  - 60a ナビゲーション情報ASP
  - 60b 音楽情報ASP
  - 60c 番組情報ASP
  - 60d 電話情報ASP

【書類名】 図面

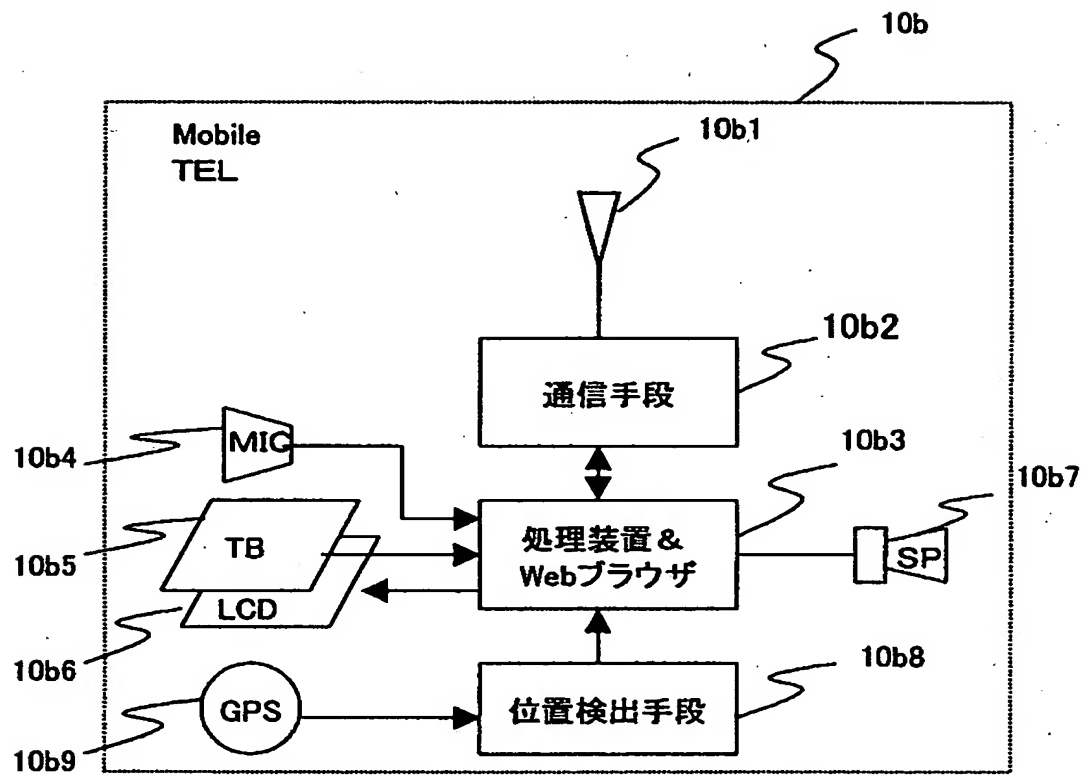
【図 1】



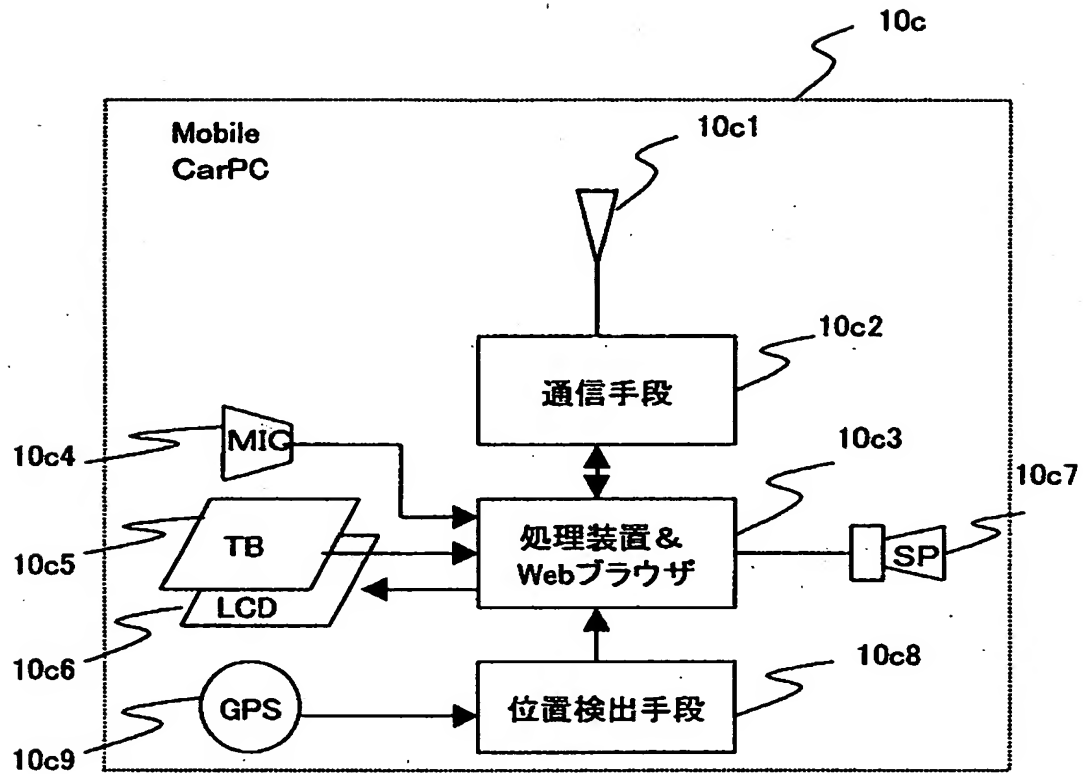
【図 2】



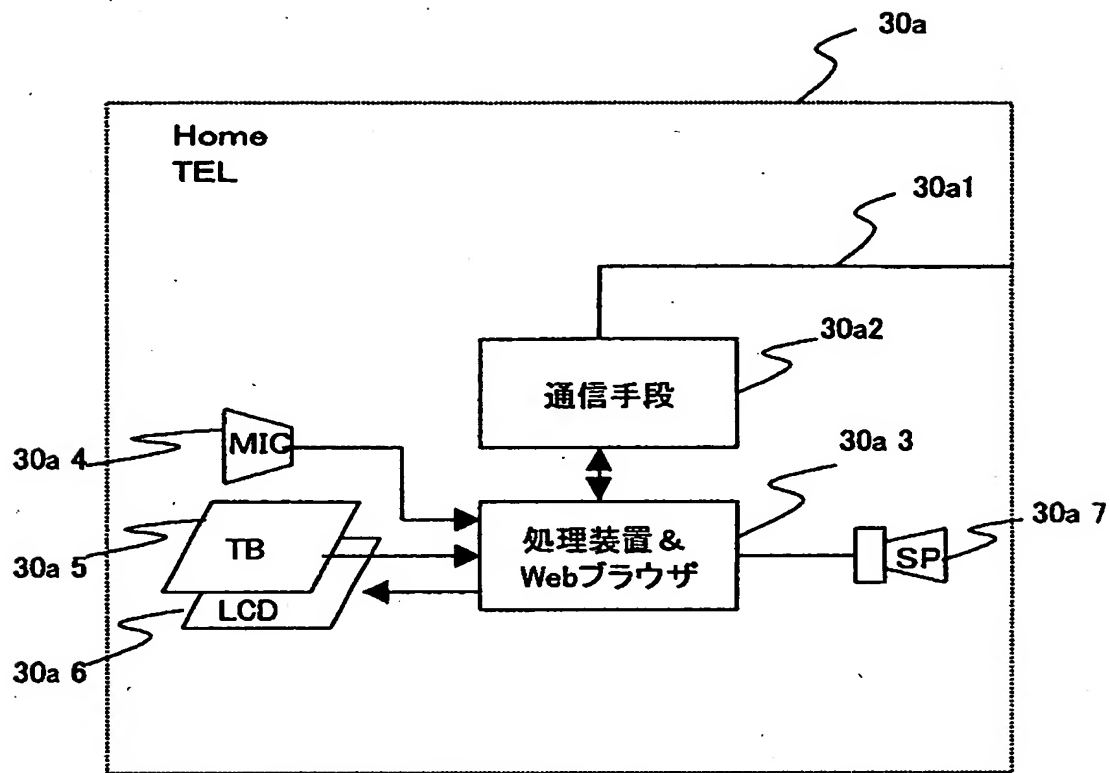
【図 3】



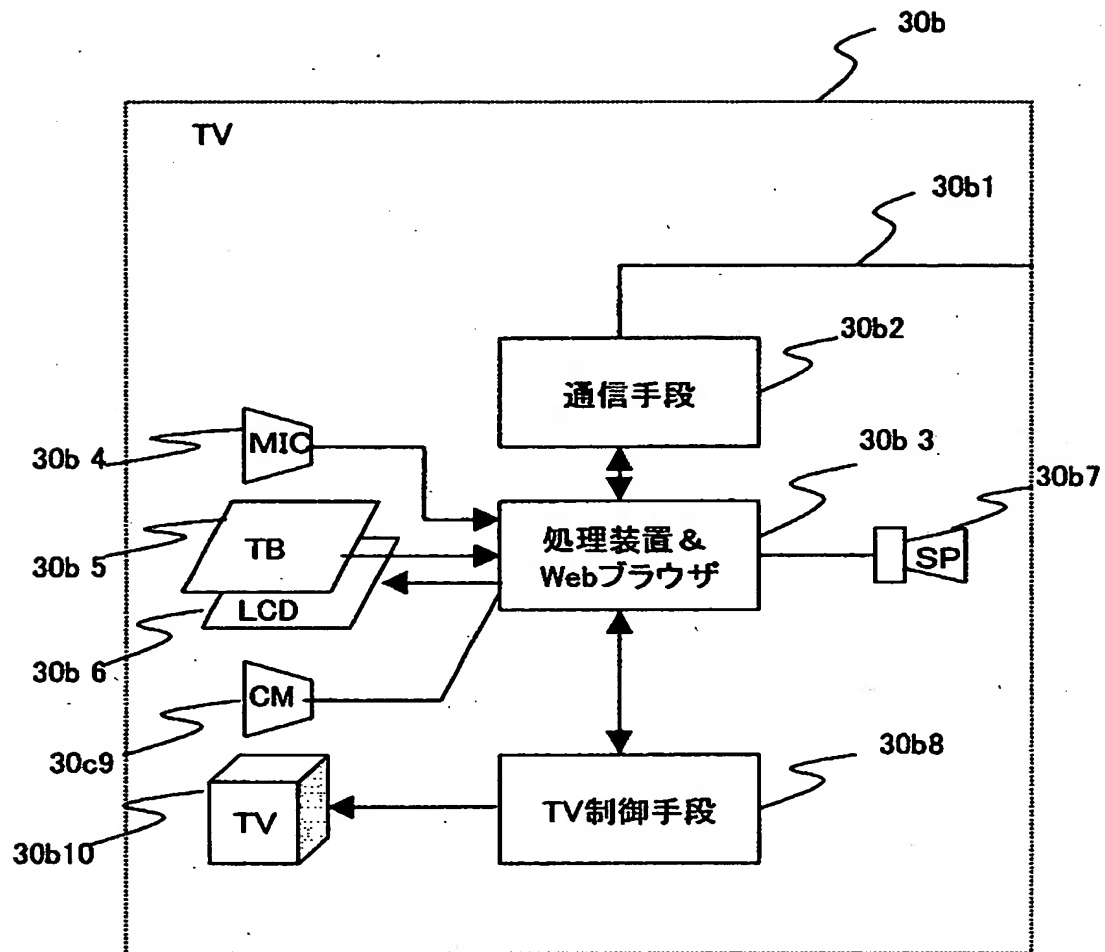
【図 4】



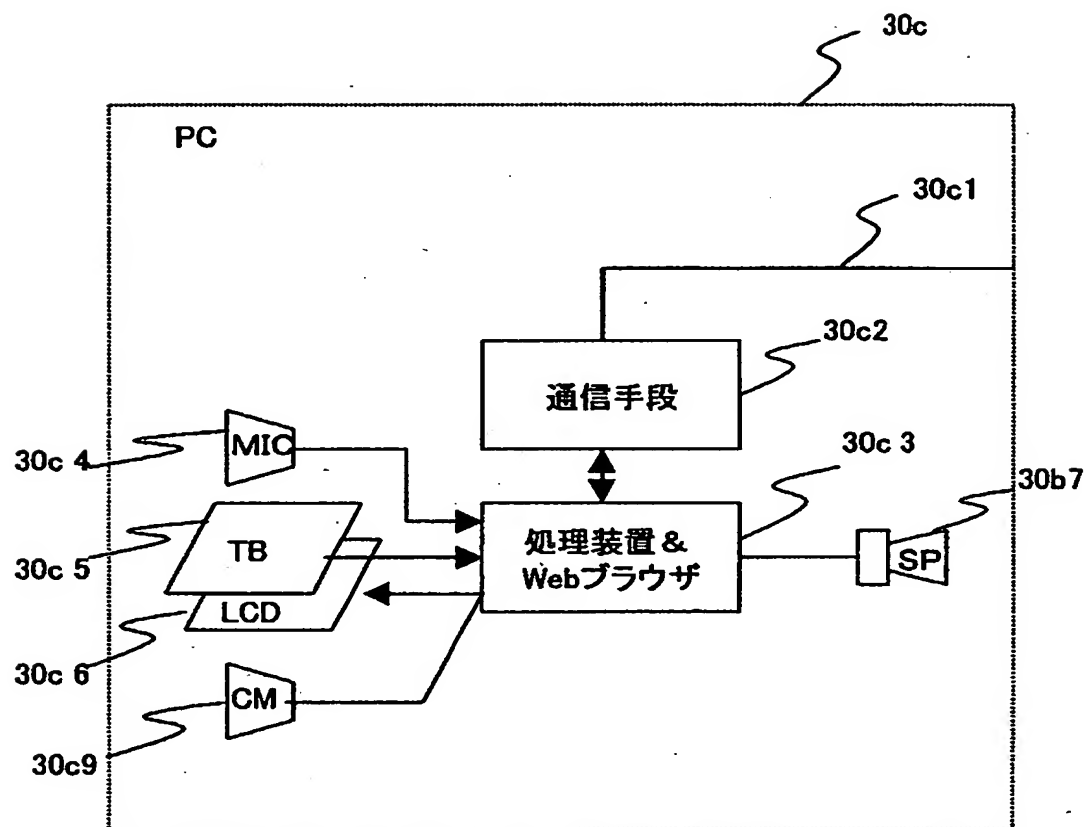
【図 5】



【図6】

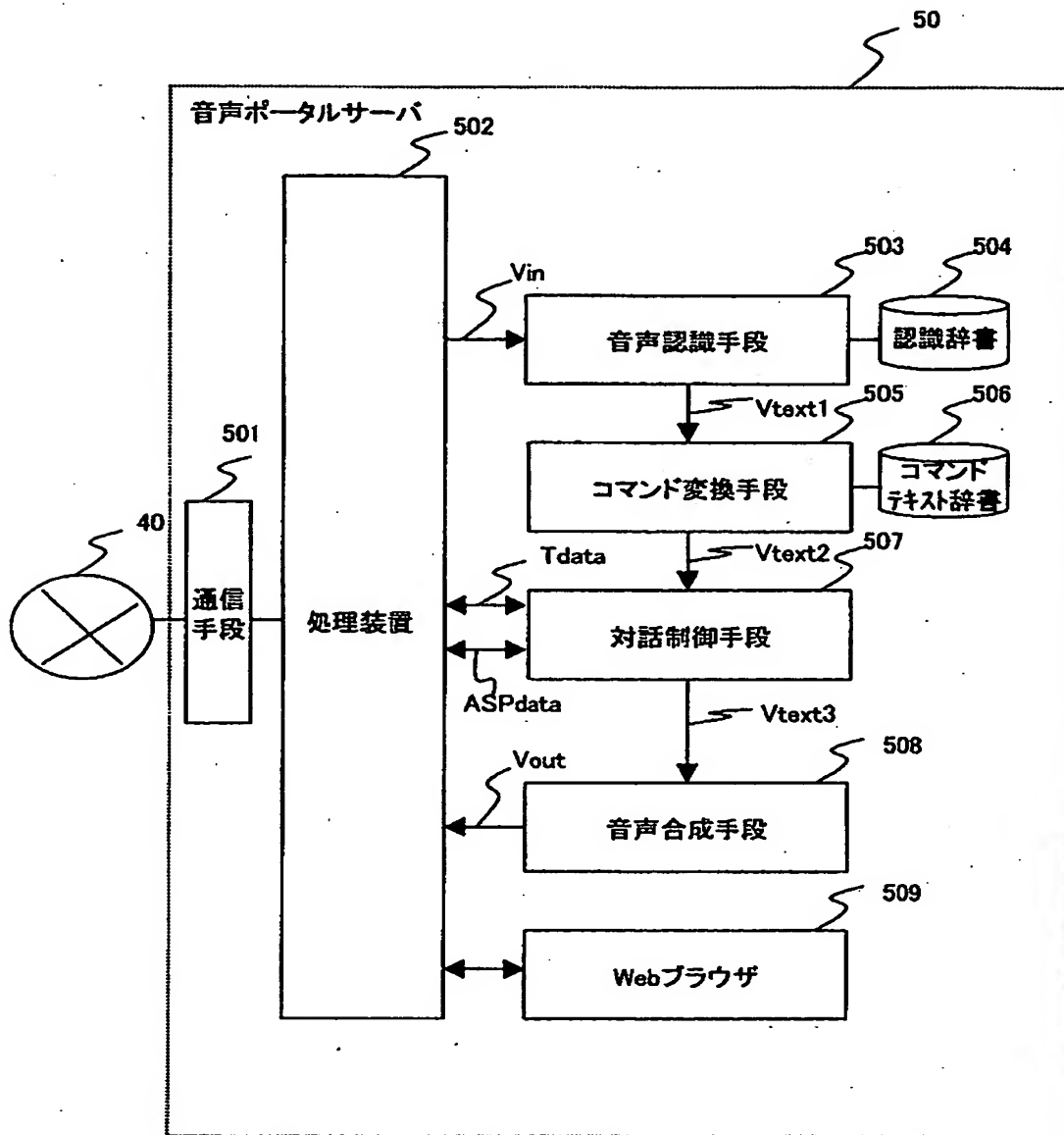


【図 7】

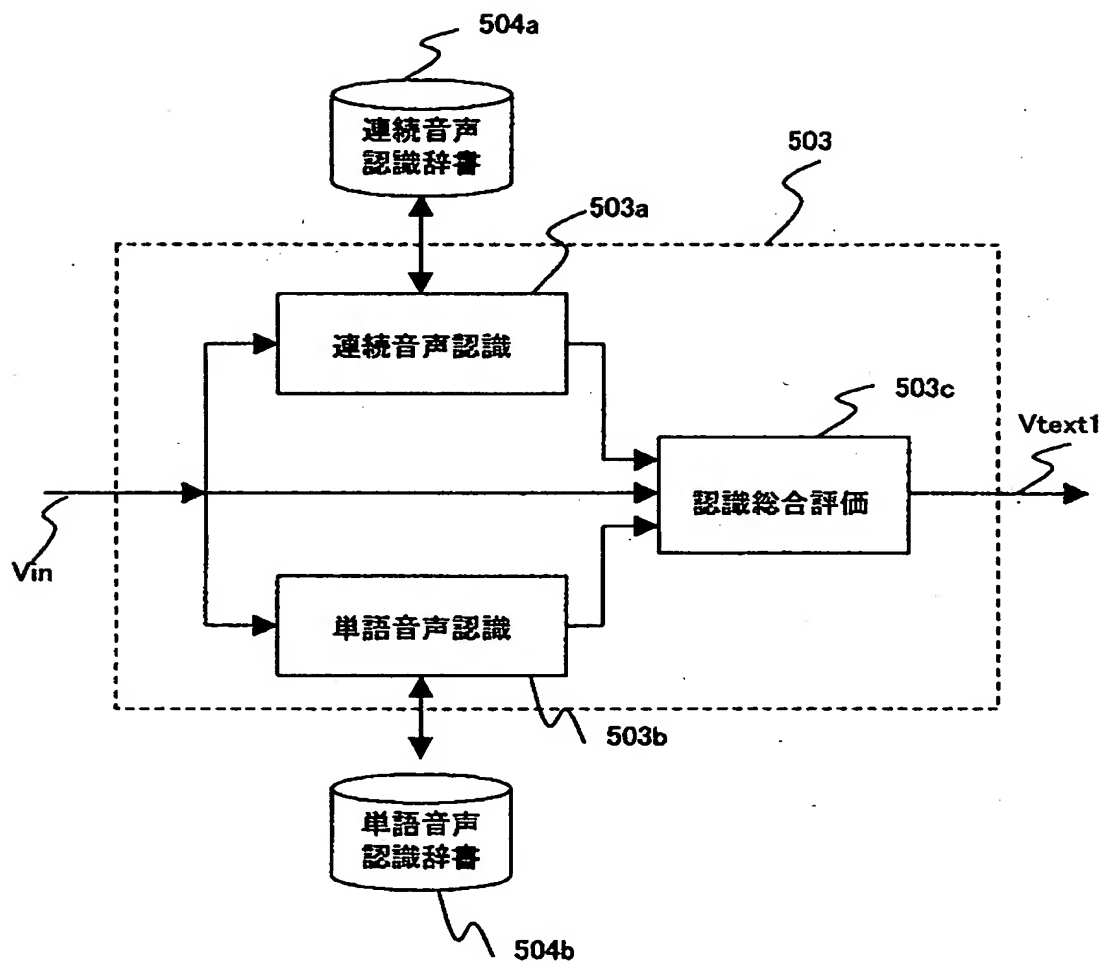




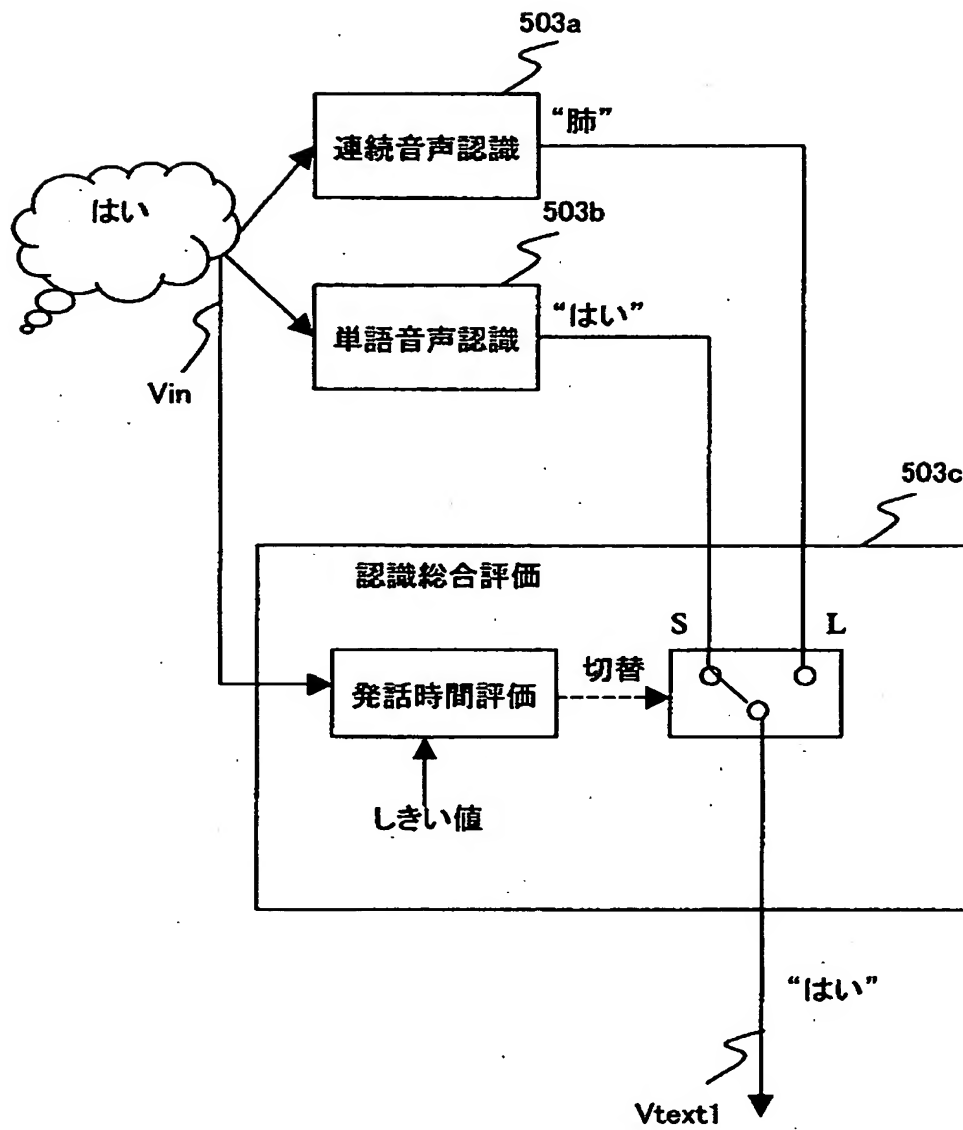
【図 8】



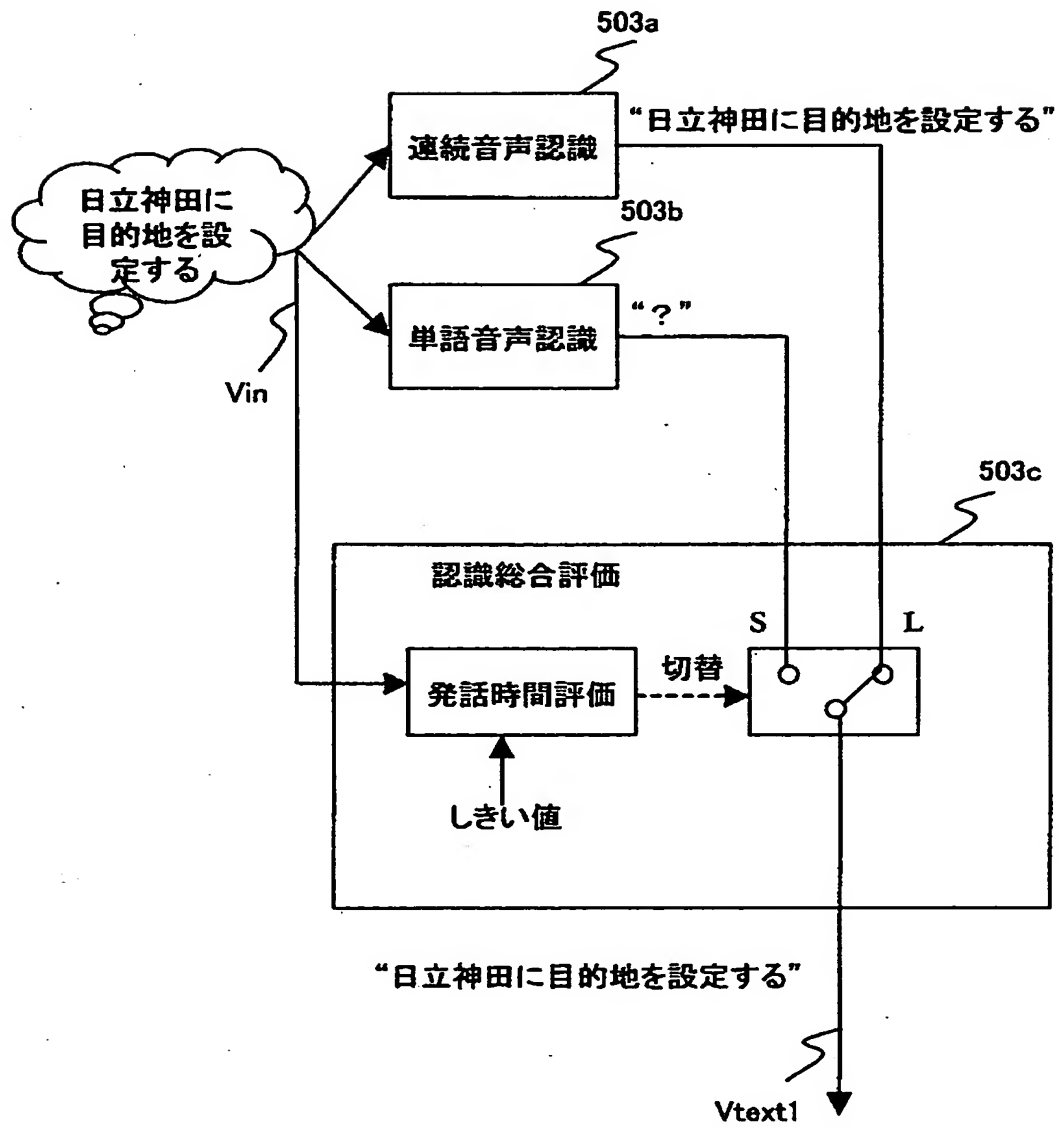
【図9】



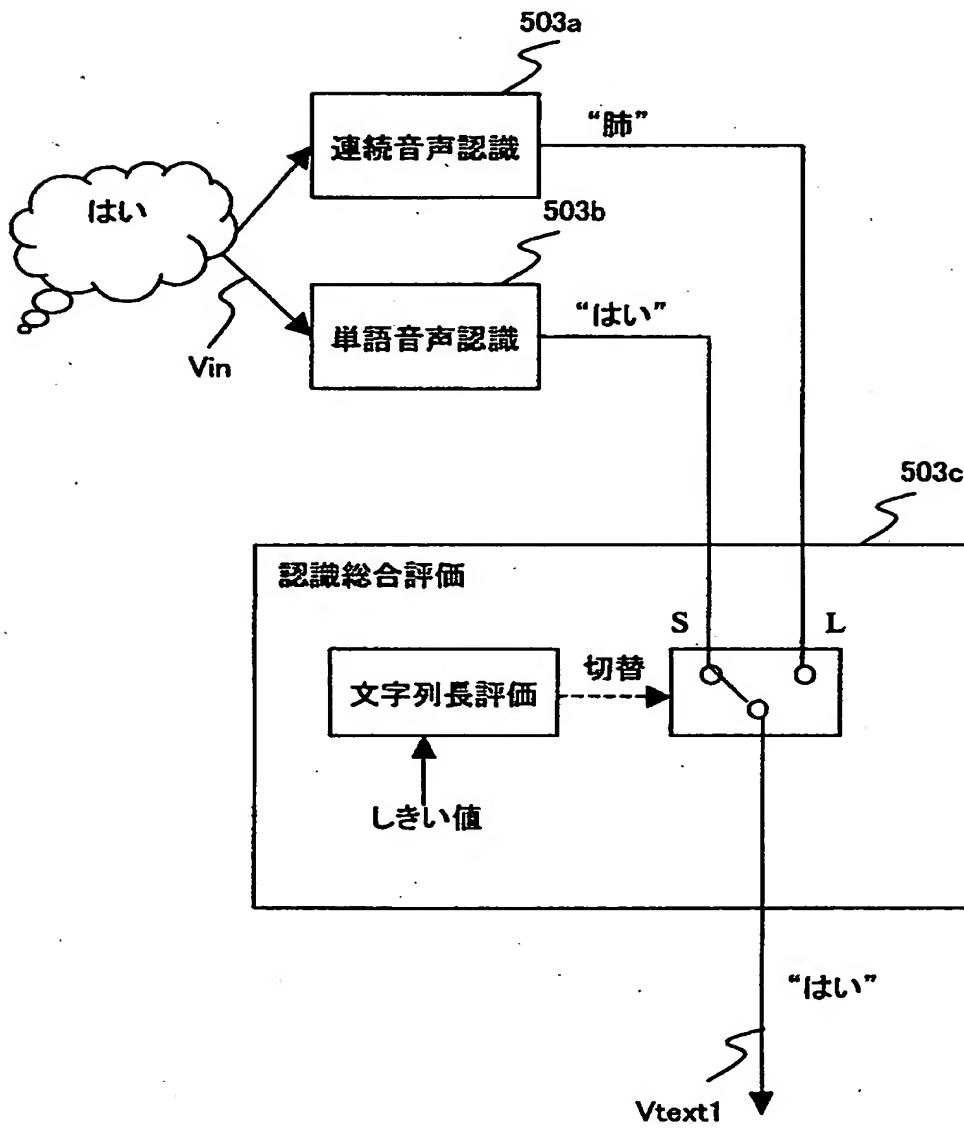
【図 10】



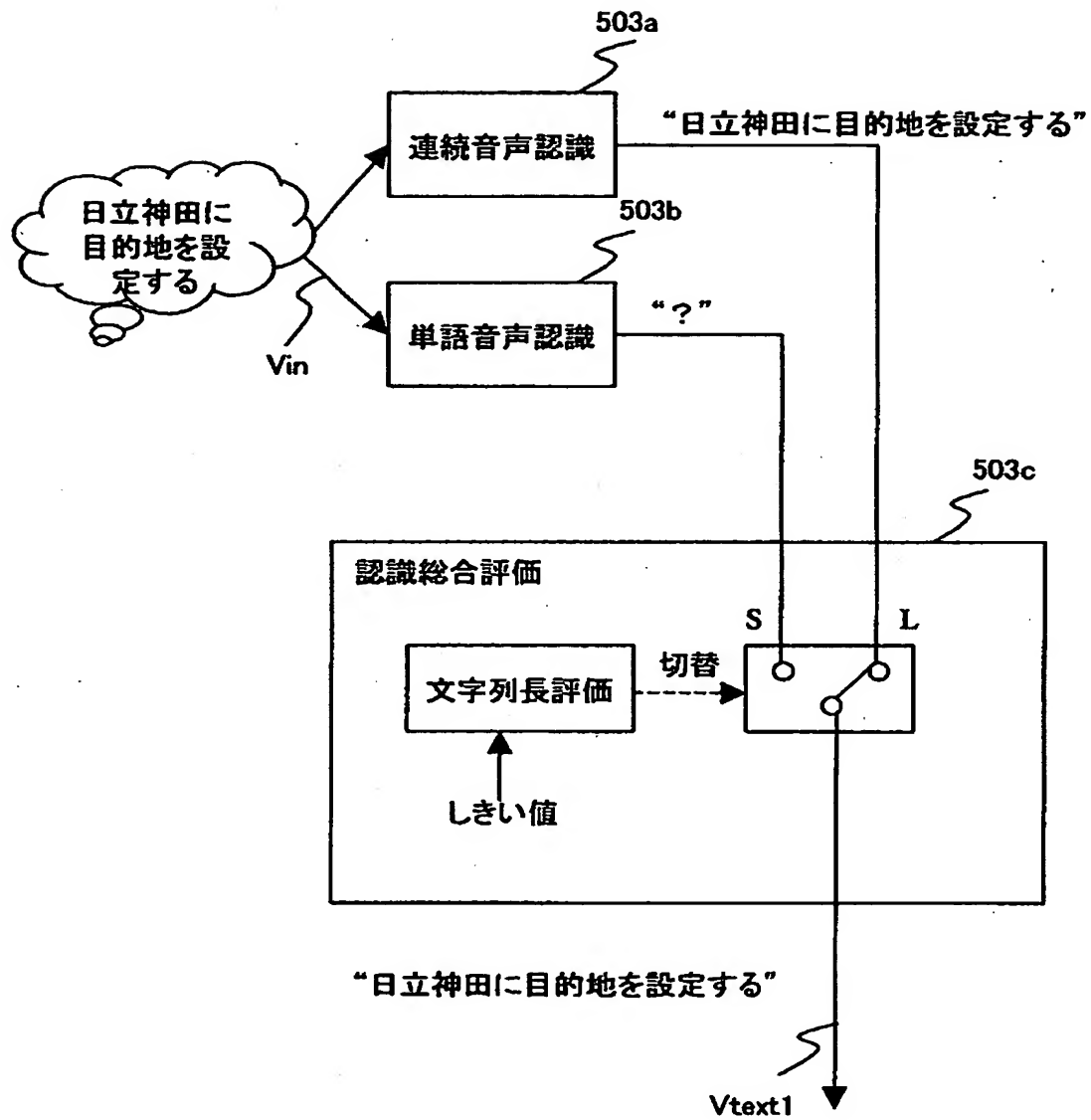
【図 11】



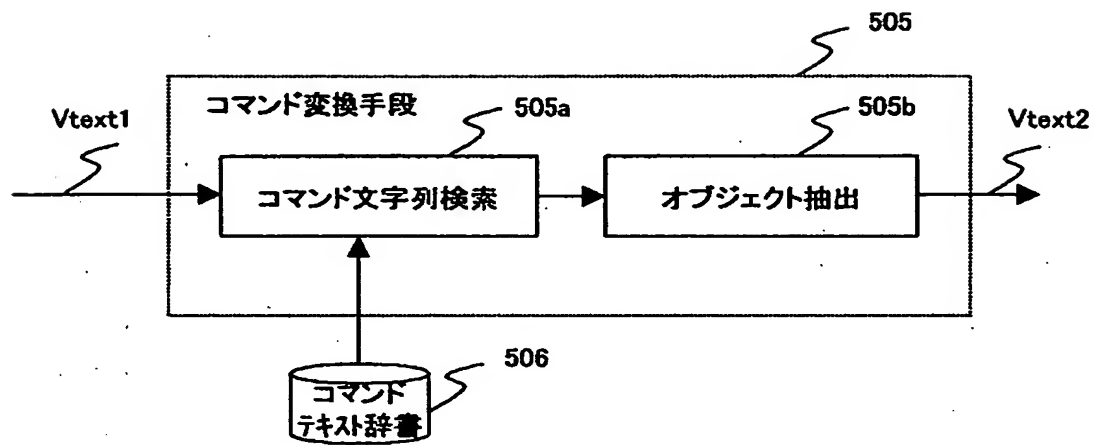
【図 12】



【図 13】



【図 14】

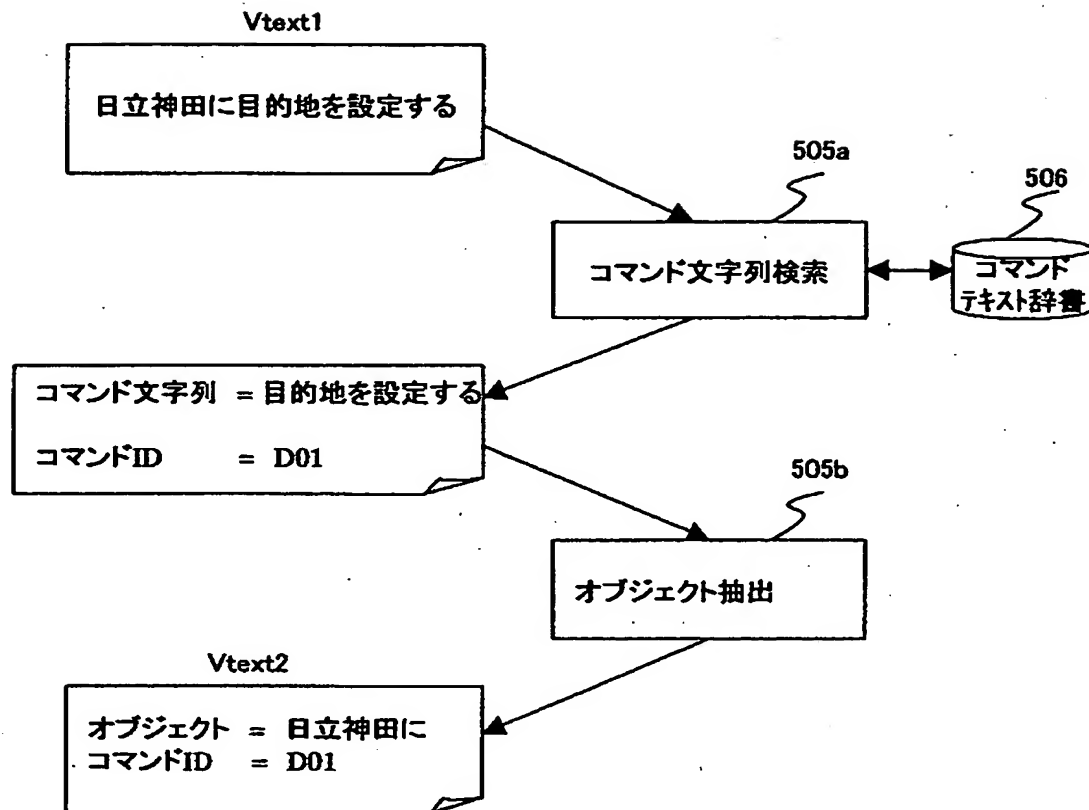


【図 15】

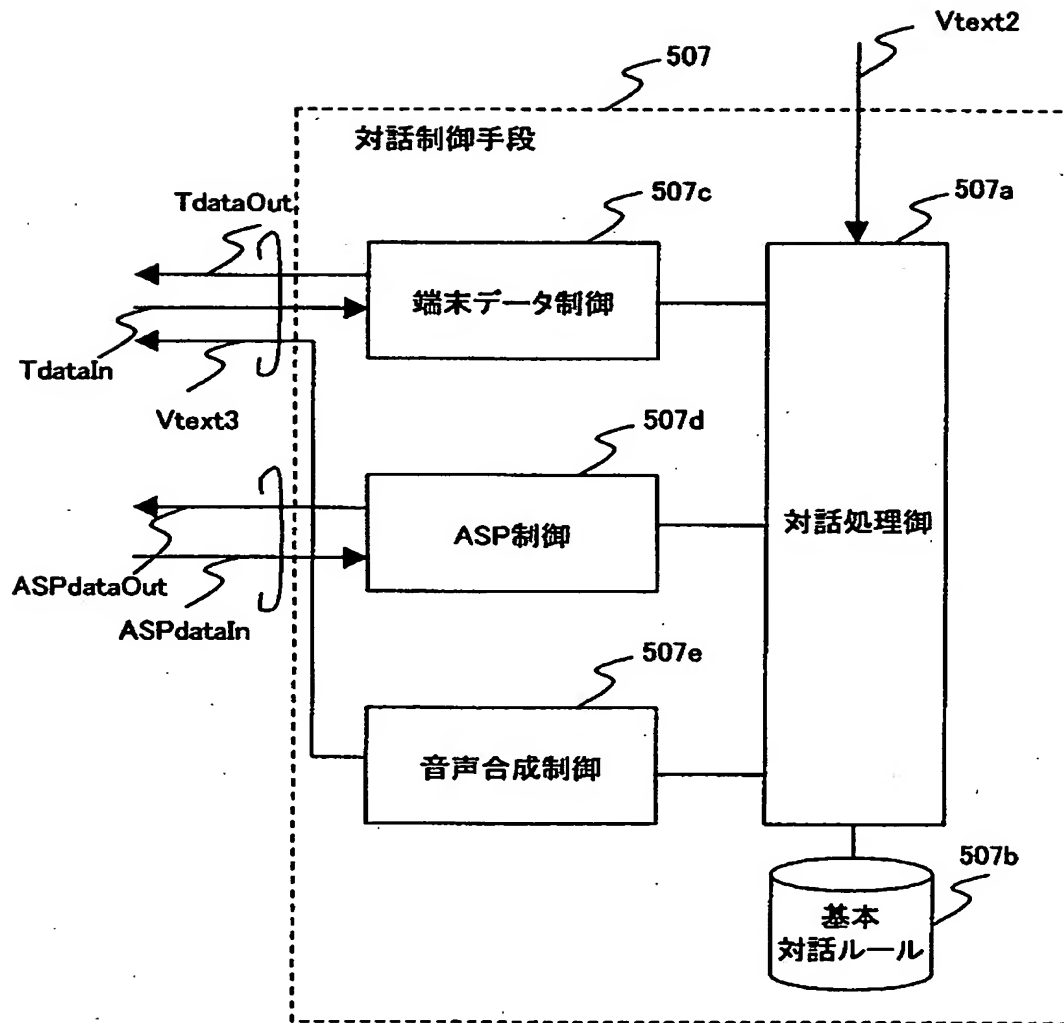
| NO | コマンドID | コマンド名1  | コマンド名2   | コマンド名3   | コマンド名4 | コマンド名5 |
|----|--------|---------|----------|----------|--------|--------|
| 1  | D01    | 目的地設定   | 目的地を設定する | 行き先を設定する | 行きたい   | 行く     |
| 2  | D02    | 目的地検索   | 目的地を探して  | 目的地を探せ   |        |        |
| 3  | D03    | 経由地設定   | 経由地を設定する | 寄りたい     | 経由したい  | 寄り道したい |
| 4  | D04    | 現在地表示   | 現在地を教える  | 現在地はどこ   |        |        |
| 5  | M01    | 音楽検索    | 曲をかけたい   | 曲を聴きたい   | 曲を探して  | 曲を探せ   |
| 6  | T01    | 番組検索    | 番組をみたい   | 番組を探して   | 番組を探せ  |        |
| 7  | T02    | 番組予約    | 番組を予約    |          |        |        |
| 8  | P01    | 電話検索    | 電話をかけて   | 電話をしたい   | 電話をして  | 話をしたい  |
| 9  | YES    | はい      | そうです     | イエス      |        |        |
| 10 | NO     | いいえ     | ちがいます    | ノー       |        |        |
| 11 | SU     | 上にスクロール | 上に移動     |          |        |        |
| 12 | SD     | 下にスクロール | 下に移動     |          |        |        |
| 13 | SR     | 右にスクロール | 右に移動     |          |        |        |
| 14 | SL     | 左にスクロール | 左に移動     |          |        |        |
| 15 | STOP   | ストップ    | 停止       |          |        |        |
| 16 | BACK   | 前へ      | 前に       |          |        |        |
| 17 | NEXT   | 次へ      | 次に       |          |        |        |
| 18 | NO1    | 1番      | 1番目      |          |        |        |
| 19 | NO2    | 2番      | 2番目      |          |        |        |
| 20 | NO3    | 3番      | 3番目      |          |        |        |
| 21 | NO4    | 4番      | 4番目      |          |        |        |
| 22 | NO5    | 5番      | 5番目      |          |        |        |



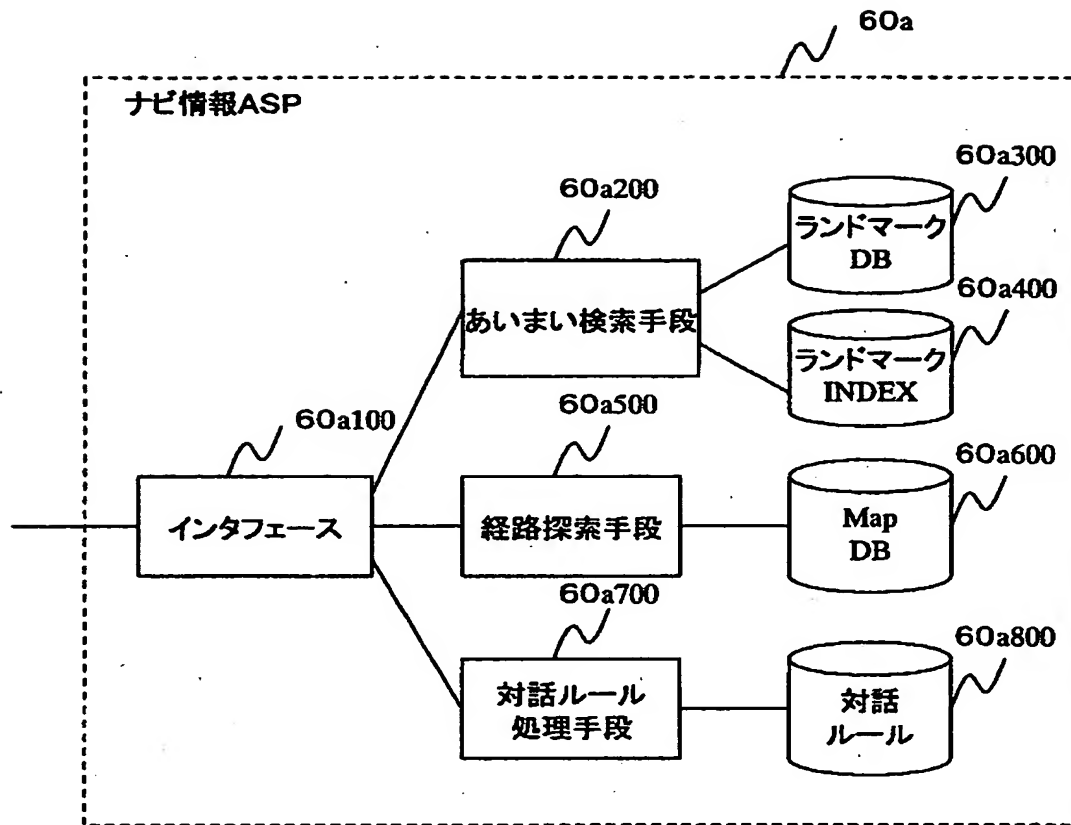
【図 16】



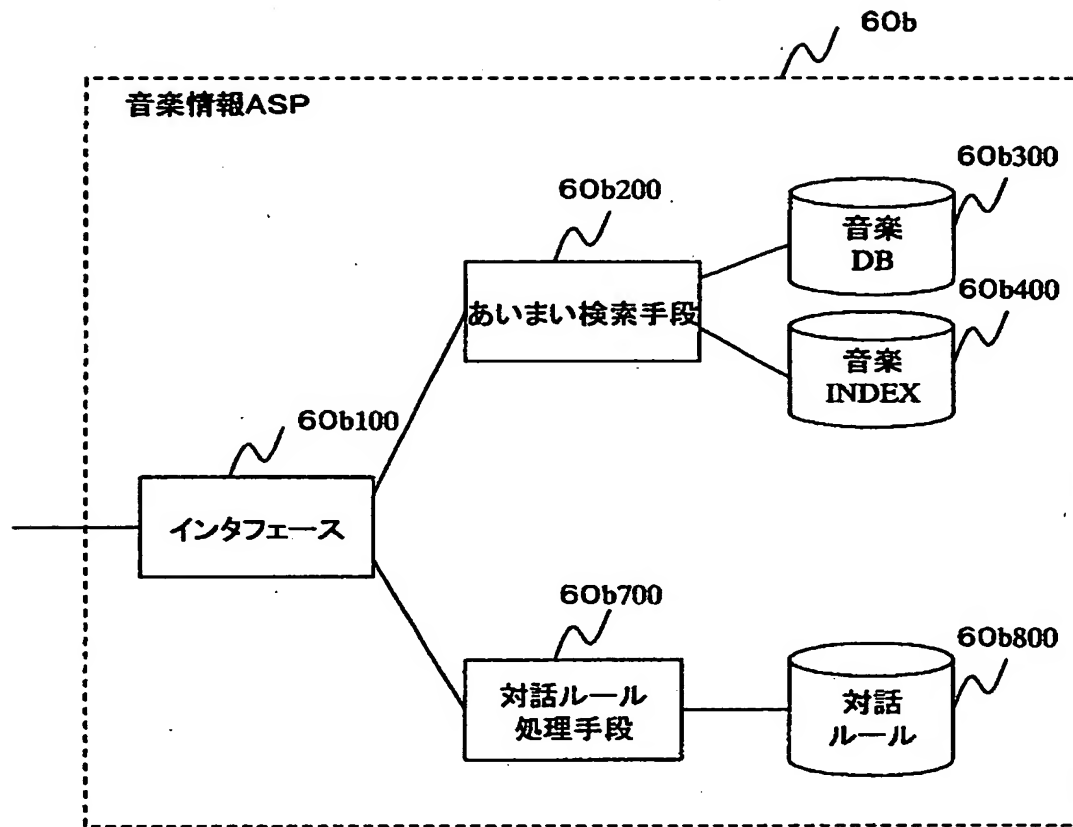
【図 17】



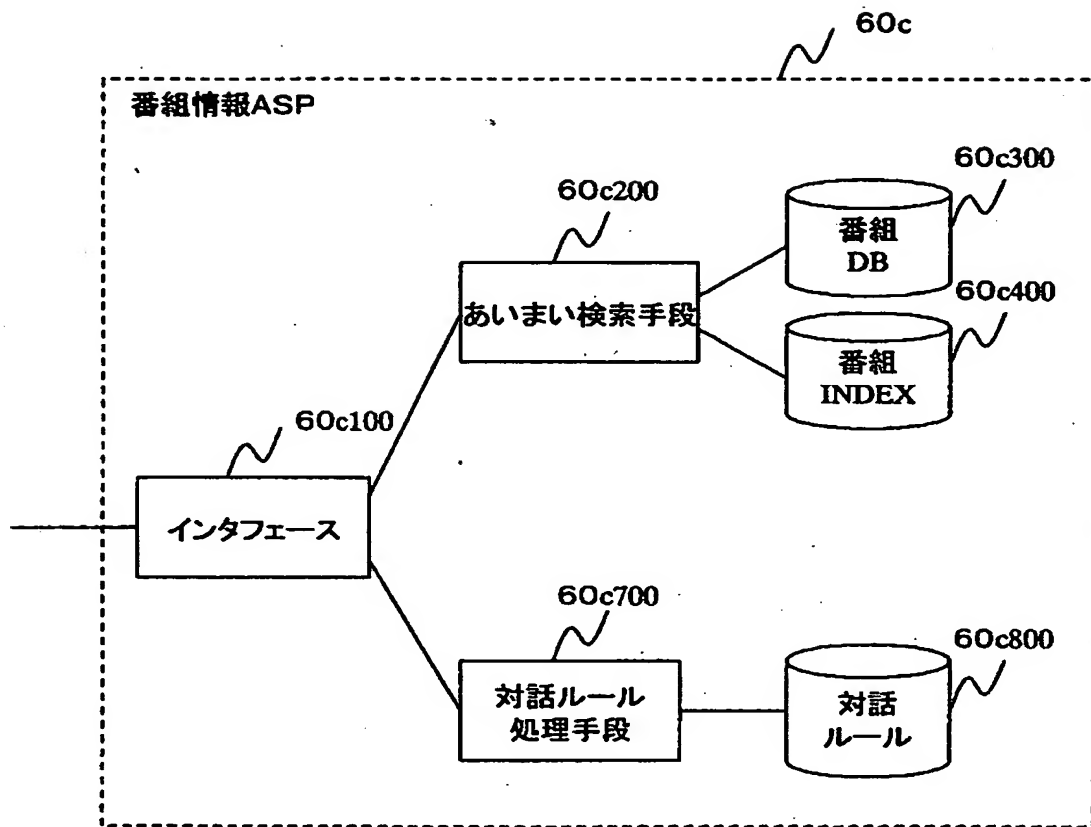
【図18】



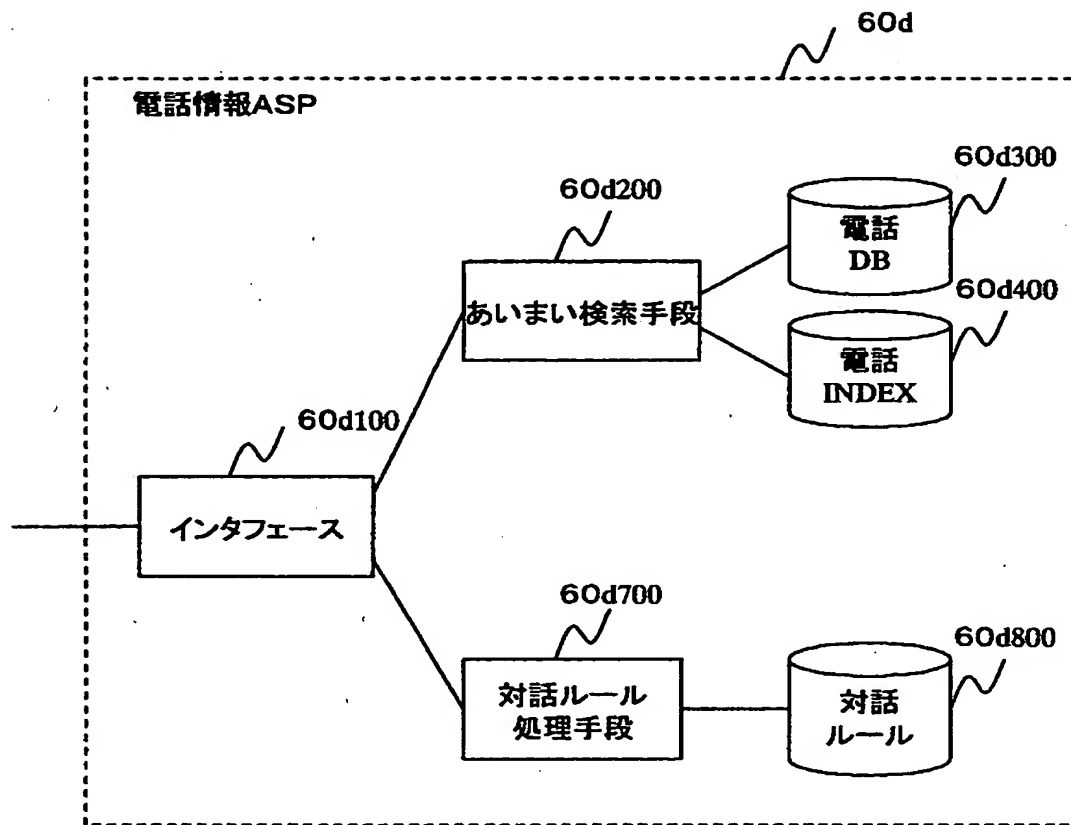
【図19】



【図20】



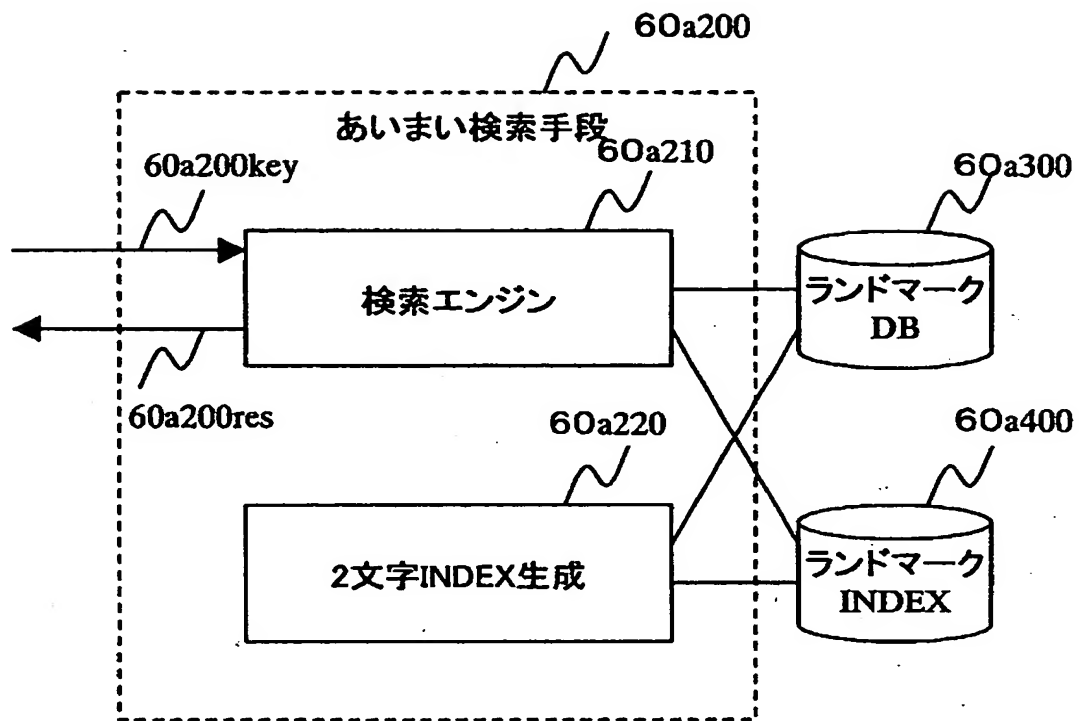
【図 2 1】



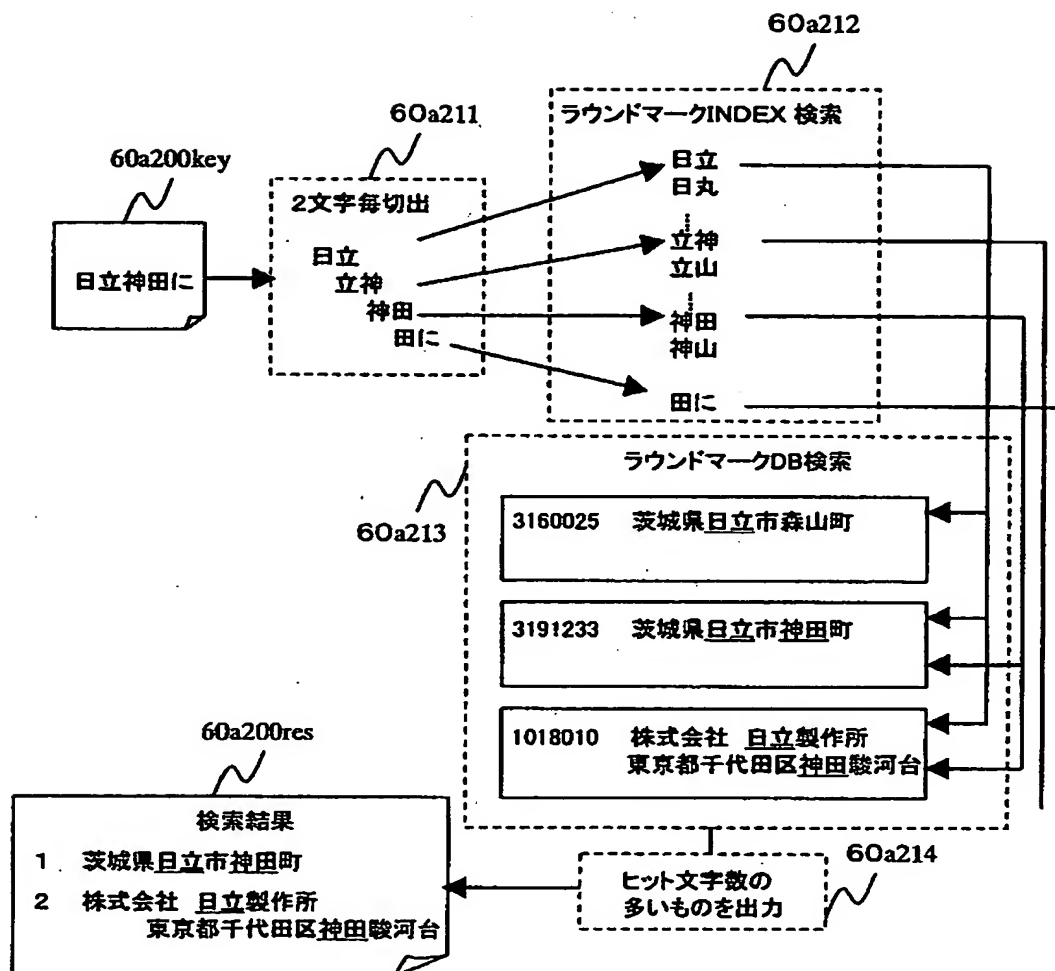
【図 2 2】

| メニュー   |       |       |       |
|--|-------|-------|-------|
| <p>■下記メニューから音声入力ください。</p> <p>■直接音声入力もできます。</p> <p style="margin-left: 20px;">(目的地設定例: 日立神田に目的地を設定する)</p> <p style="margin-left: 20px;">(電話検索例: 小泉一郎さんに電話をかけたい)</p> |       |       |       |
| 目的地設定  | 目的地検索 | 経由地設定 | 現在地表示 |
| <div style="text-align: center; margin-top: 10px;">音楽検索</div>  |       |       |       |
| 番組検索   | 番組予約  |       |       |
| <div style="text-align: center; margin-top: 10px;">電話検索</div>  |       |       |       |

【図 2 3】

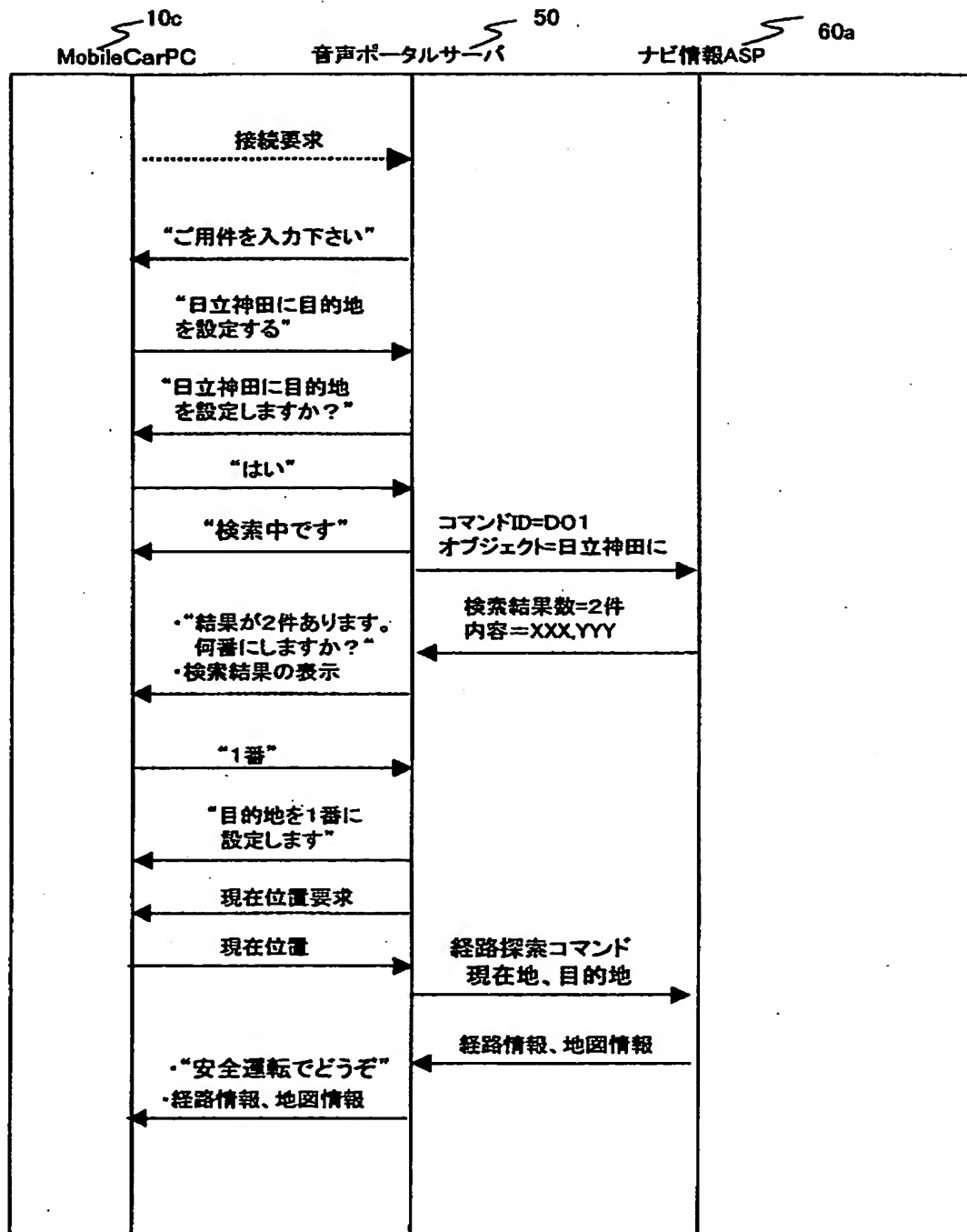


【図24】

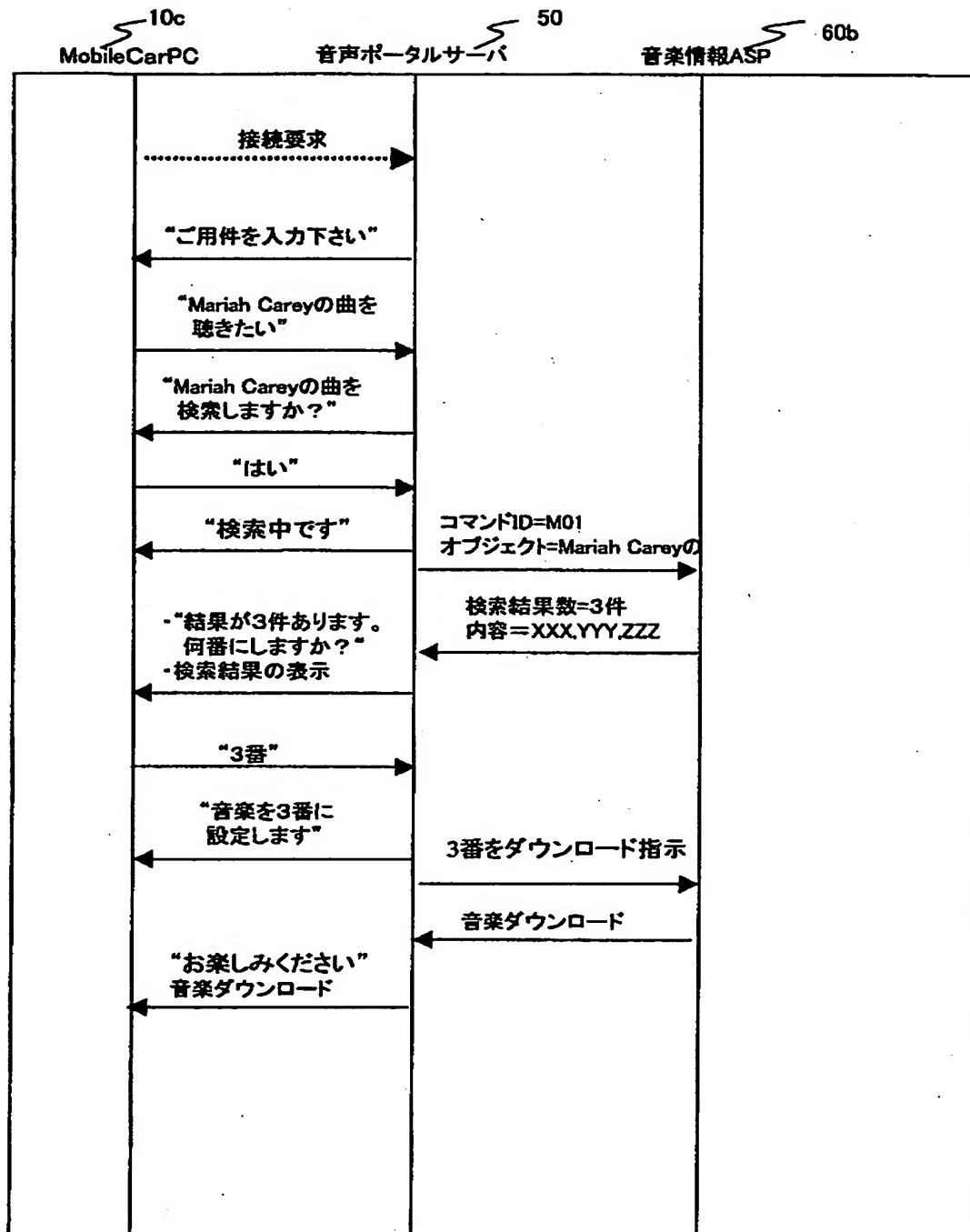




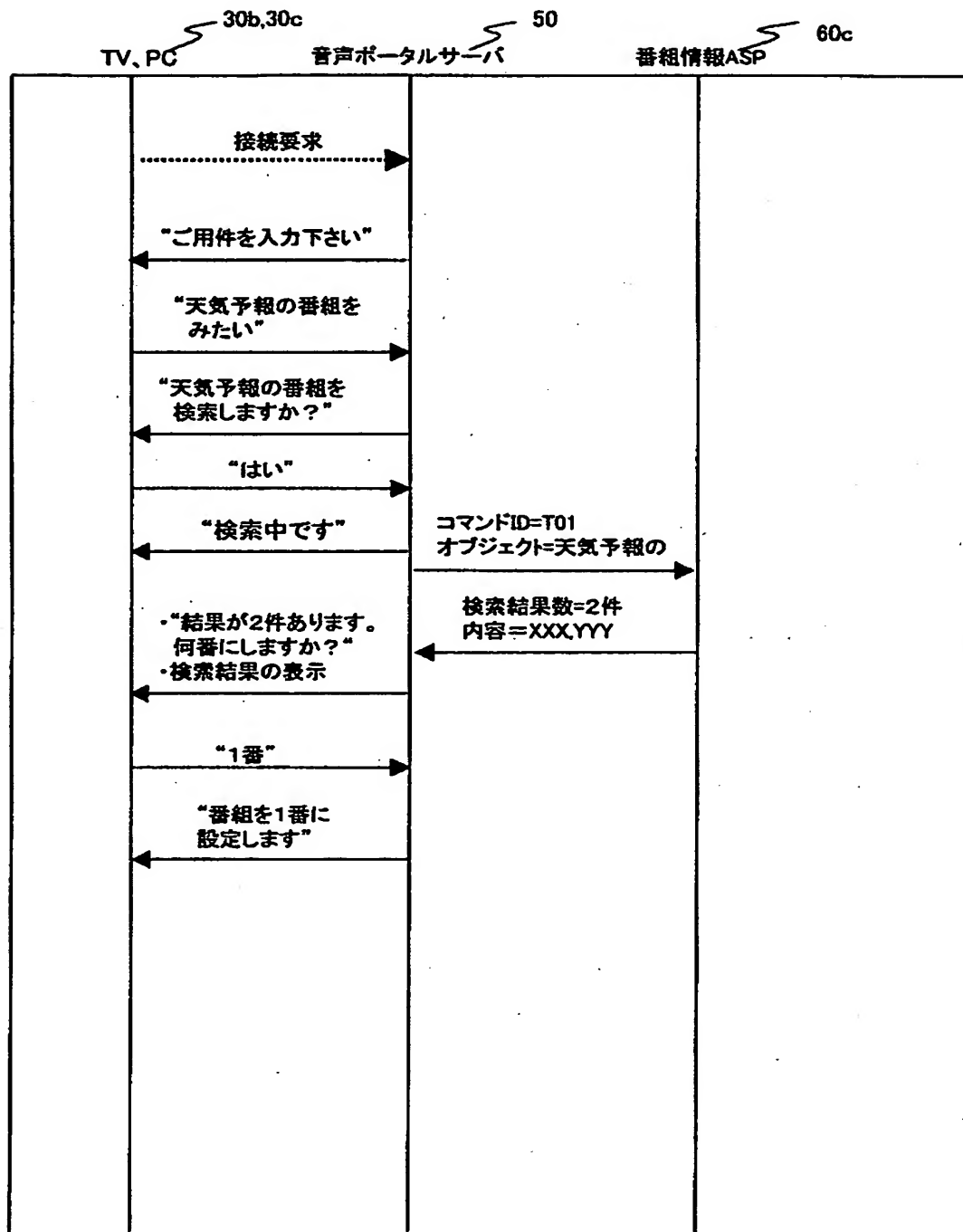
【図 25】



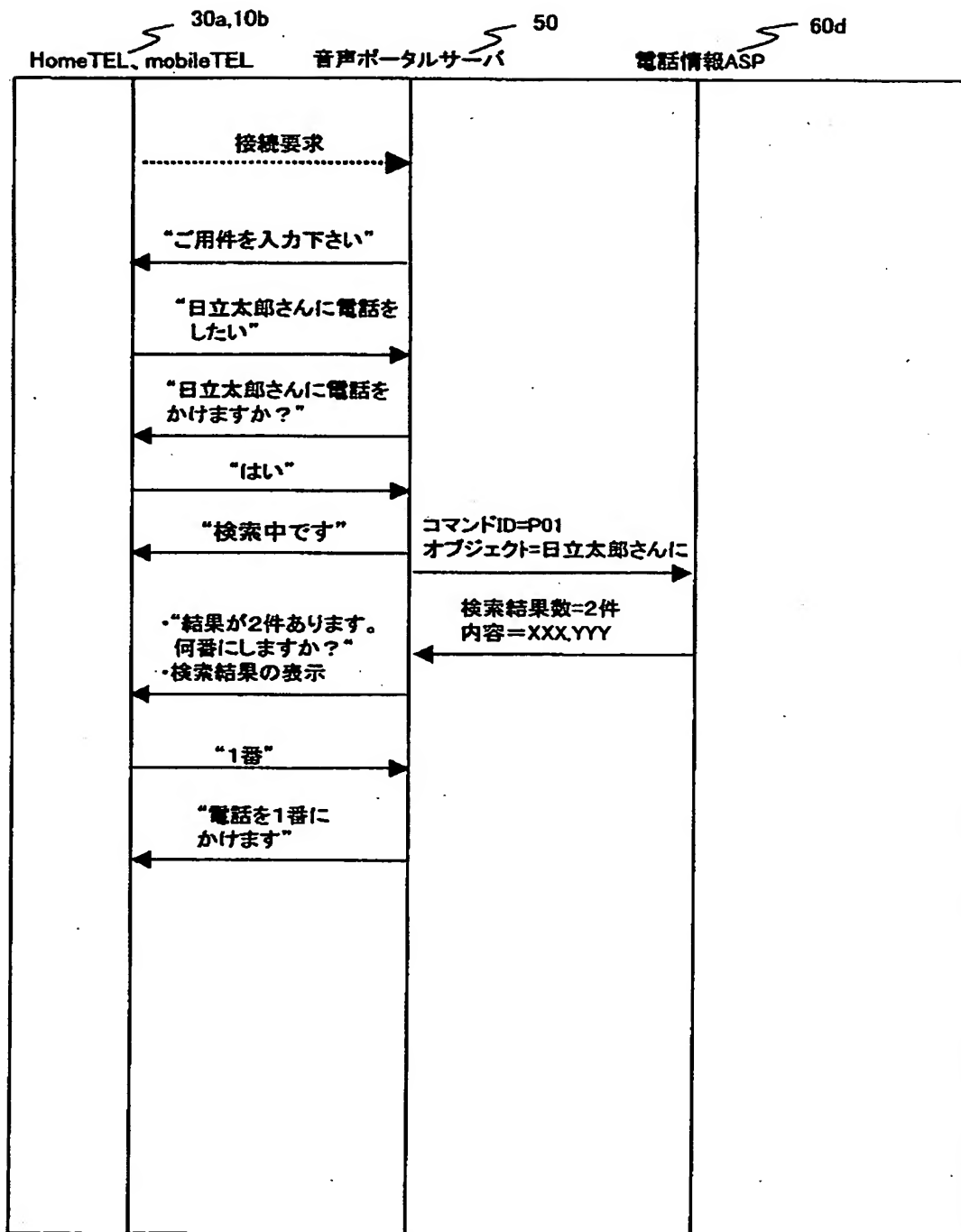
【図 26】



【図 27】



【図 28】



【書類名】            要約書

【要約】

【課題】    PDA、携帯電話などの移動端末や家庭電話、TV、PCなどの固定端末から音声でネットにアクセスし、地図情報、音楽情報、番組情報、電話情報などを提供するプロバイダからサービスを受ける音声入力システムを提供する。

【解決手段】    音声入出力手段とアクセス状況表示手段とを備えた音声入力端末10、30と、音声を取り込みテキストとして認識する音声認識手段と認識されたテキストをコマンドテキスト辞書と照合しコマンドテキストとオブジェクトテキストとに分離するコマンド変換手段と分離されたテキストに基づいて各種情報を提供するプロバイダにアクセスしてサービスを受けこのサービスを音声入力端末に出力する対話制御手段とを備えた音声ポータルサーバ50と、音声ポータルサーバから受信したコマンドテキストおよびオブジェクトテキストに基づき情報を検索しその結果を音声ポータルサーバにサービスするプロバイダ60とからなる音声入力システム。

【選択図】            図1

出 願 人 履 歴 情 報

識別番号 [000005108]

|          |                    |
|----------|--------------------|
| 1. 変更年月日 | 1990年 8月31日        |
| [変更理由]   | 新規登録               |
| 住 所      | 東京都千代田区神田駿河台4丁目6番地 |
| 氏 名      | 株式会社日立製作所          |